

NUCLEUS

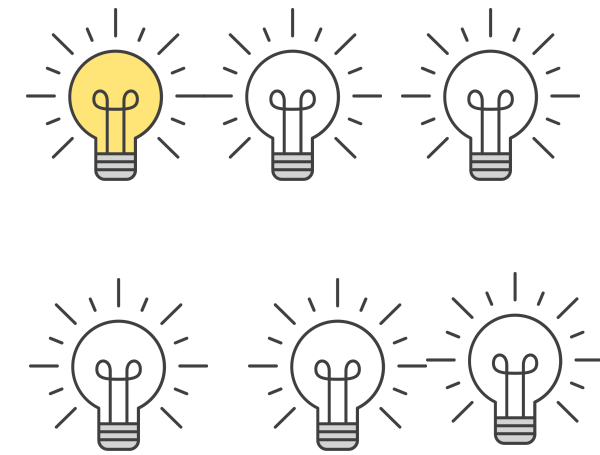
Take command of your energy use with total home monitoring.

By: Yashaswini Thokala,
Jaya Keerthi Varagani,
Nehemie Joseph,
Chandrika Rai

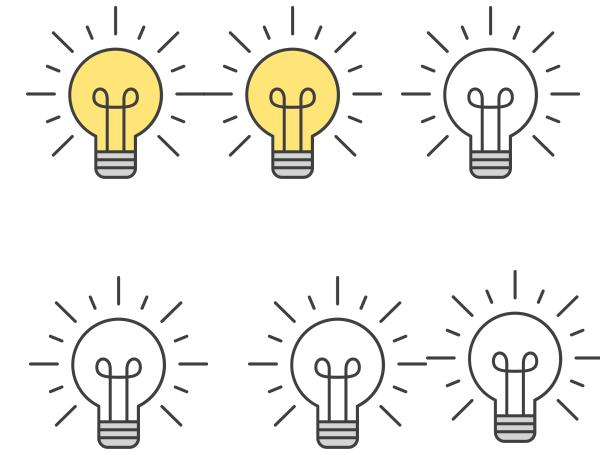
ENLIGHTENING TIMELINE

1977

The development of smart meters that predawn the birth of the internet has evolved in calculating the consumption of household items.



ENLIGHTENING TIMELINE



1977

The development of smart meters that predawn the birth of the internet has evolved in calculating the consumption of household items.

2016

The United Kingdom started a national smart meter roll-out to create a cheaper and more efficient smart grid.



ENLIGHTENING TIMELINE



1977

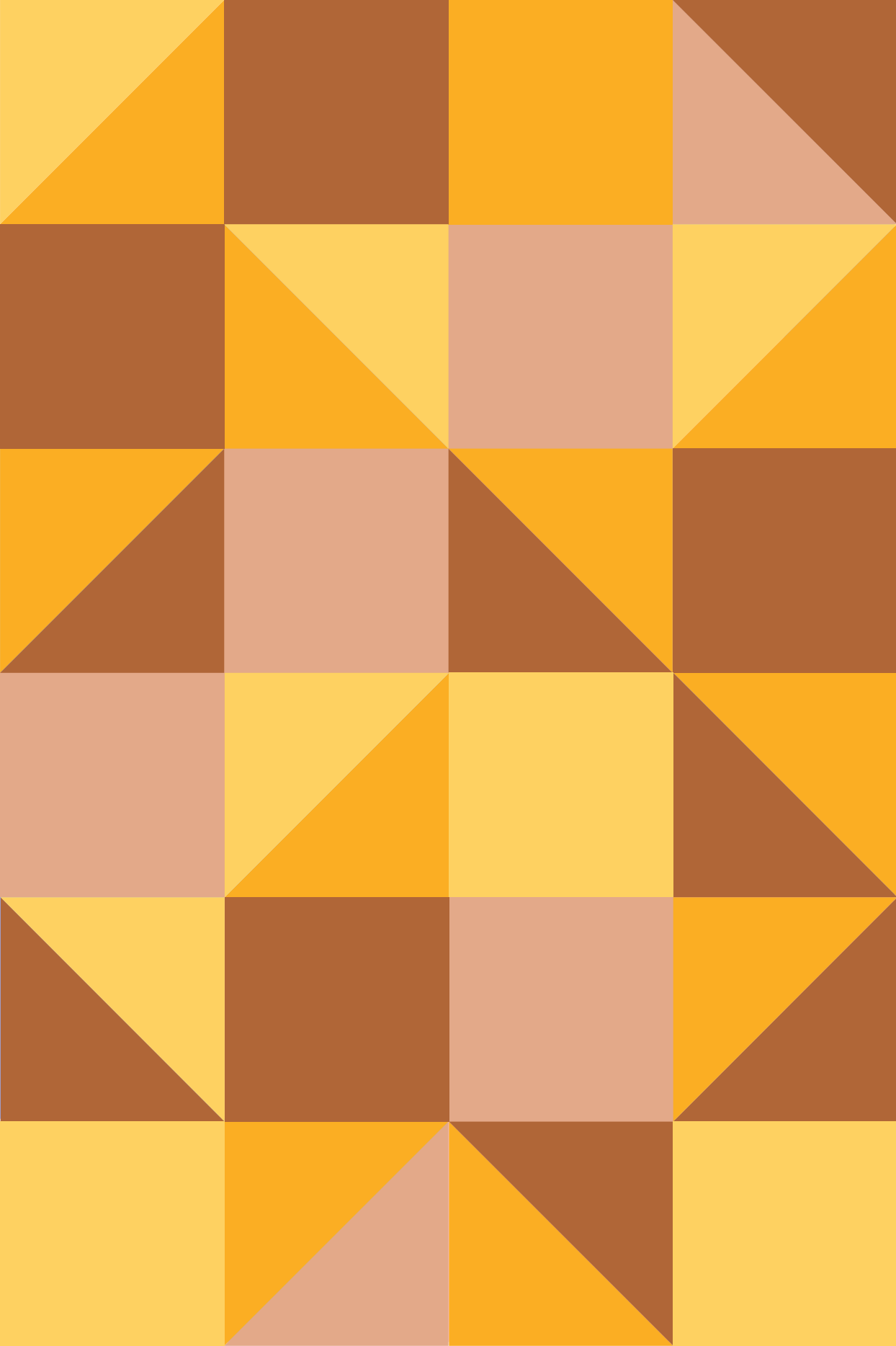
The development of smart meters that predawn the birth of the internet has evolved in calculating the consumption of household items.

2016

The United Kingdom started a national smart meter roll-out to create a cheaper and more efficient smart grid.

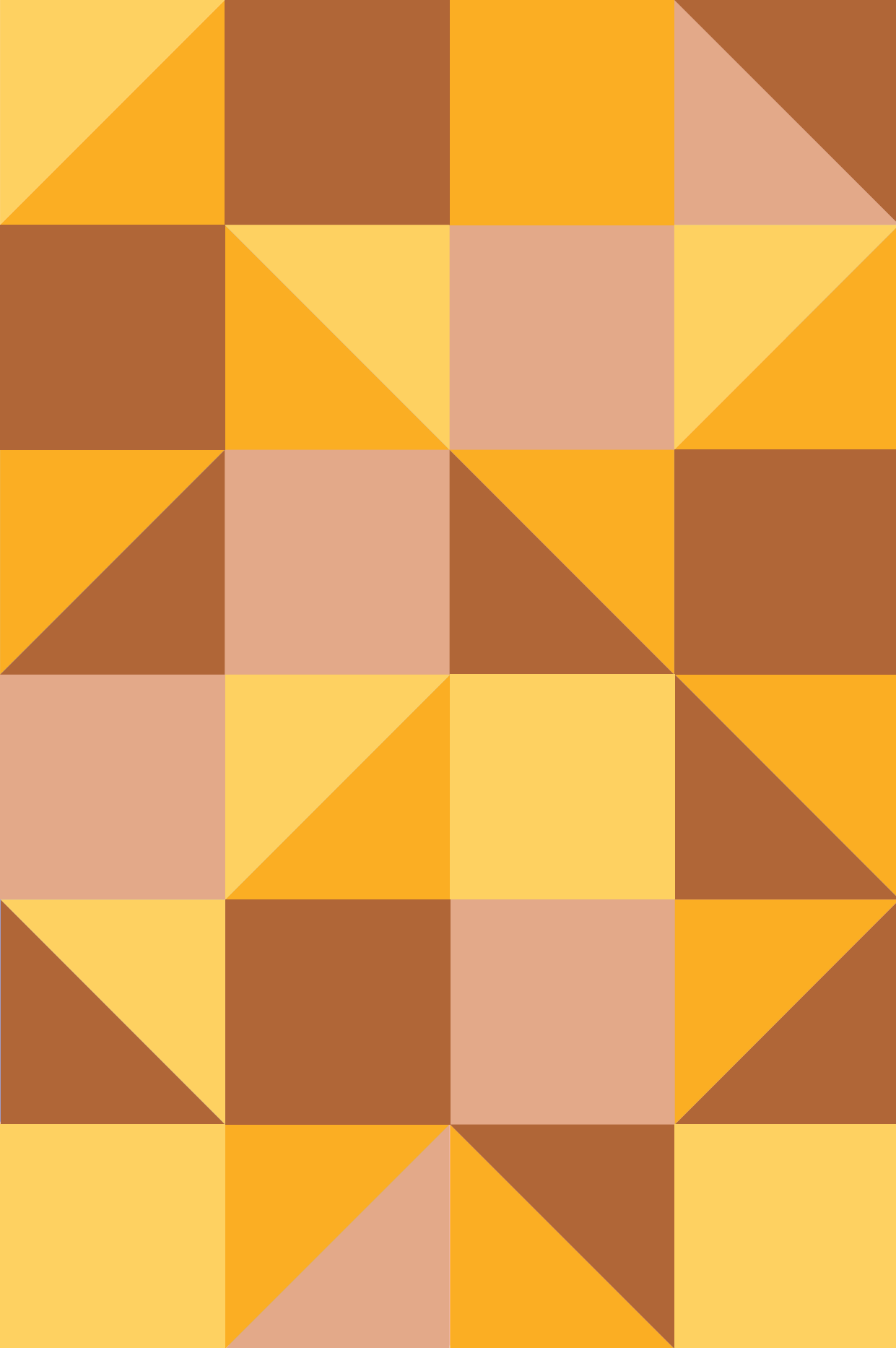
2020

The International Energy Agency, warns of a Global Energy Crisis.



ENLIGHTENING TIMELINE





ENLIGHTENING TIMELINE



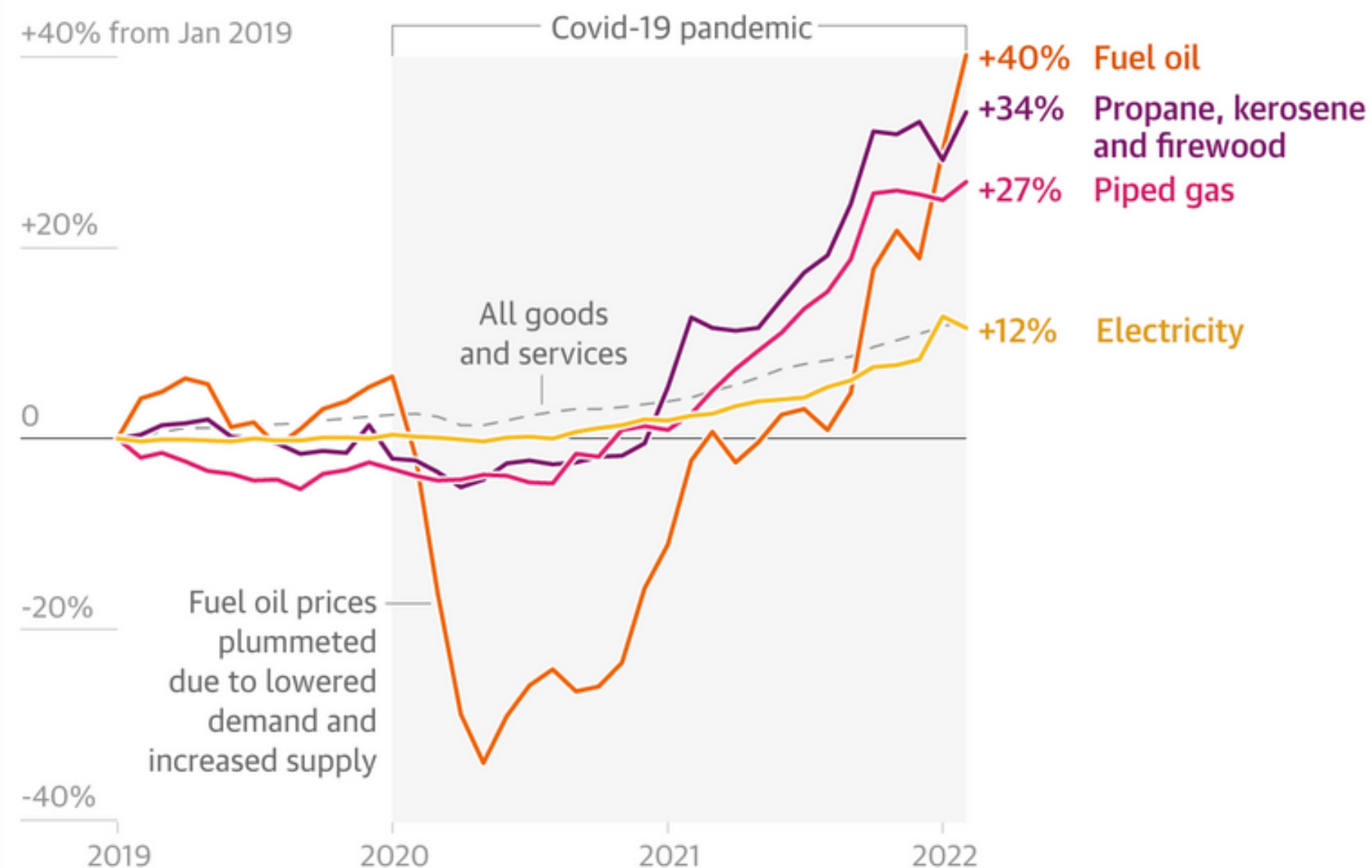
WHAT ABOUT YOU?



ENLIGHTENING TIMELINE

Home energy prices have outpaced inflation

The percent change in US prices since 2019



Guardian graphic. Source: US Bureau of Labor Statistics. Note: Monthly prices from January 2019 to February 2022, seasonally adjusted. <https://www.nytimes.com/2022/05/03/business/energy-environment/high-electric-bills-summer.html>

YOU FOCUS CONTENT PRESENTATION

01. Research Questions

What is the goal?

02. Methodologies

Source

Machine Learning Methods

03. Key Takeaways

'What about you?' Take two

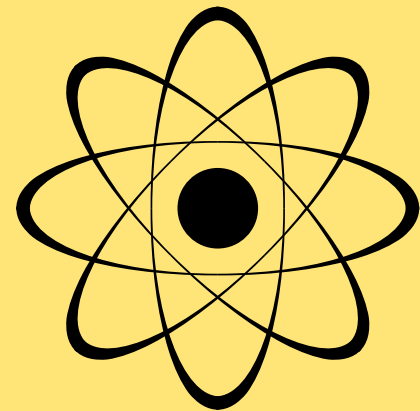
04. Improvements

What we did do wrong?

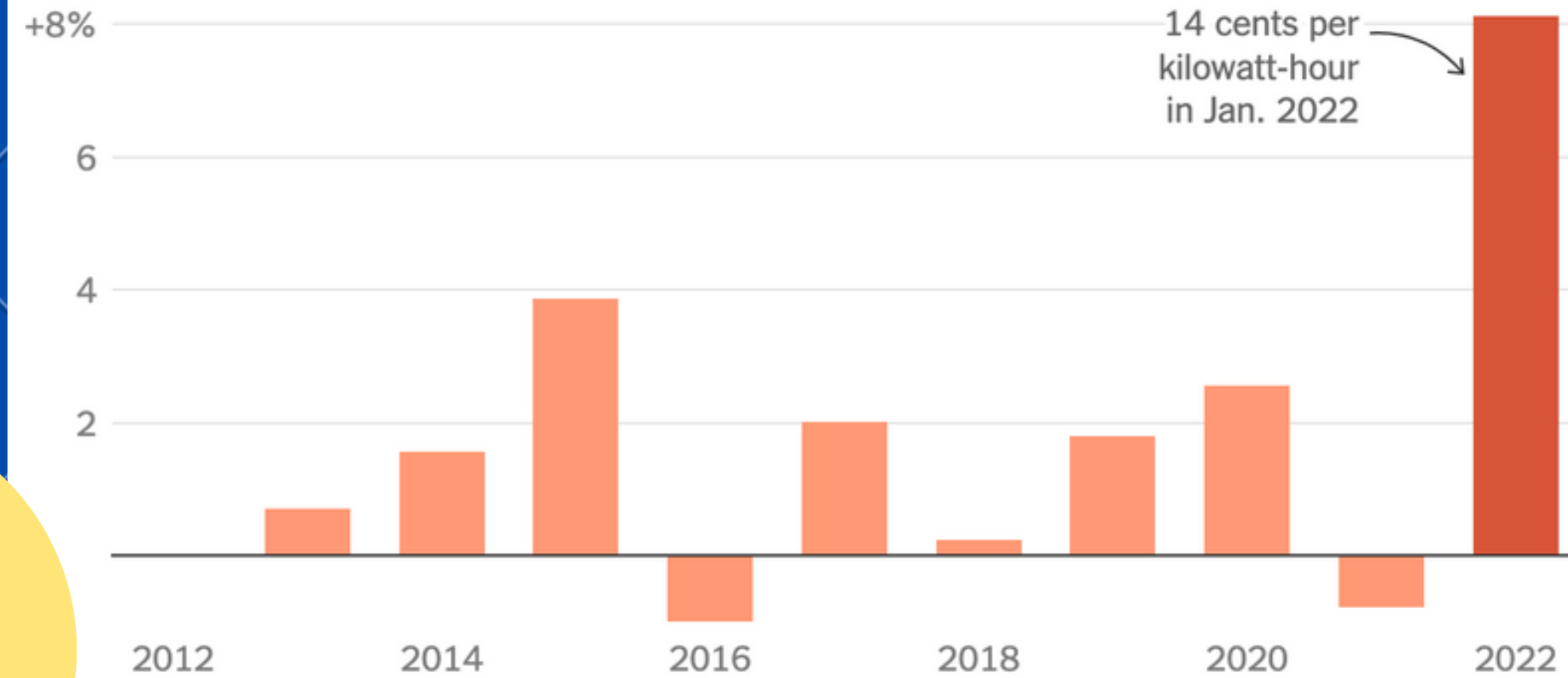
How can it be improved?



OUR NAME IS NUCLEUS.



The Rising Price of Home Electricity in the U.S.
Year-over-year change in average prices in January each year



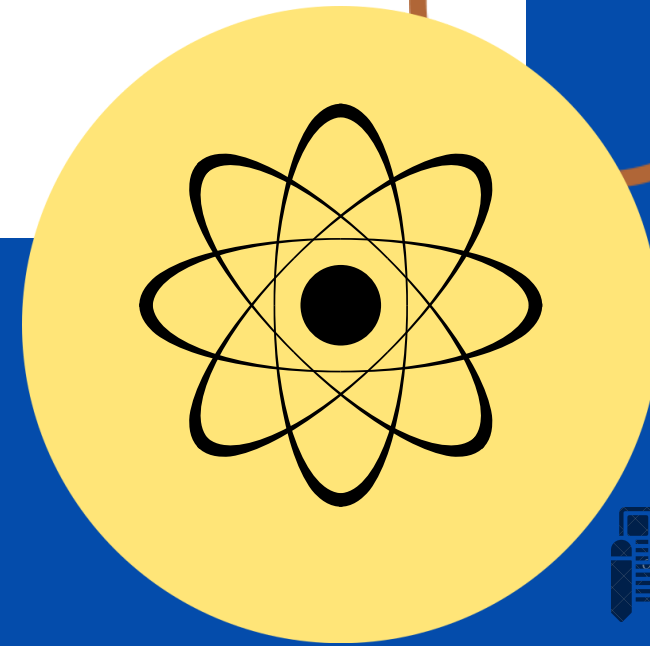
Source: U.S. Energy Information Administration • By The New York Times
<https://www.nytimes.com/2022/05/03/business/energy-environment/high-electric-bills-summer.html>



Predict residential electricity consumption using various household and weather-related factors.



**REGIONAL-
DIVISION
SIGNIFICANCE**



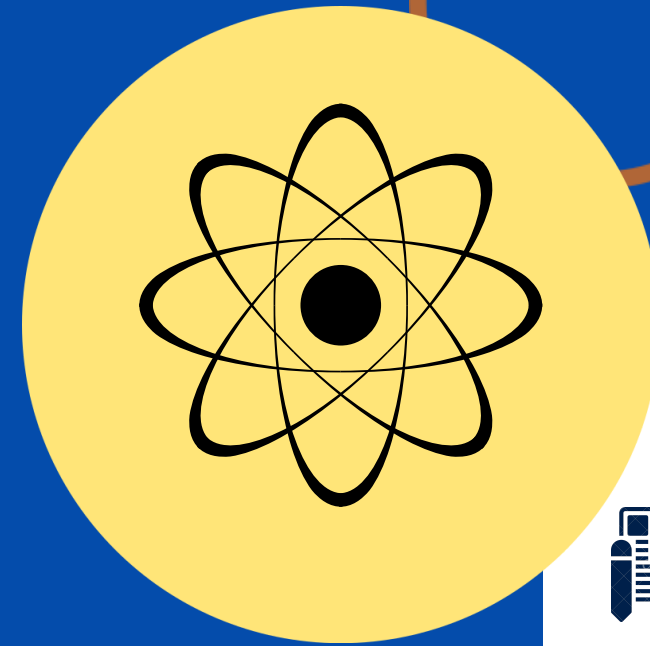
**POPULATION
CLASSIFICATIONS**

**COMMON
HOUSEHOLD
CHARACTERISTICS**



**REGIONAL-
DIVISION
SIGNIFICANCE**

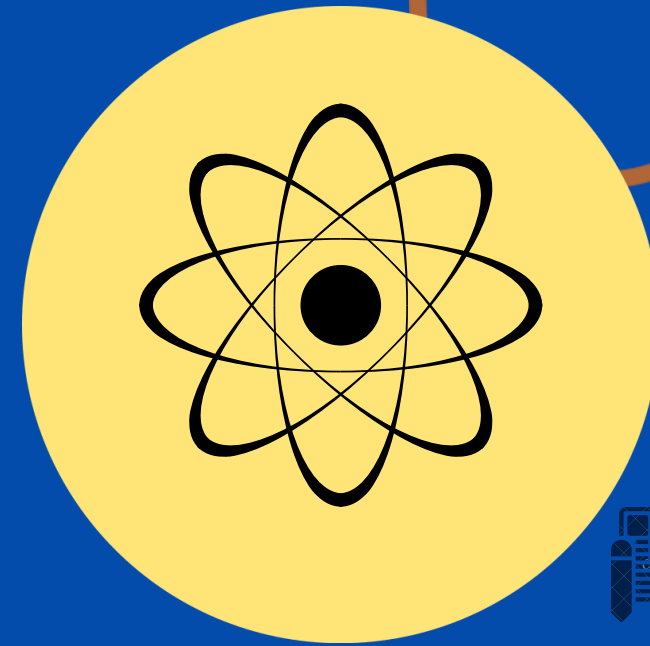
**POPULATION
CLASSIFICATIONS**



**COMMON
HOUSEHOLD
CHARACTERISTICS**



**REGIONAL-
DIVISION
SIGNIFICANCE**



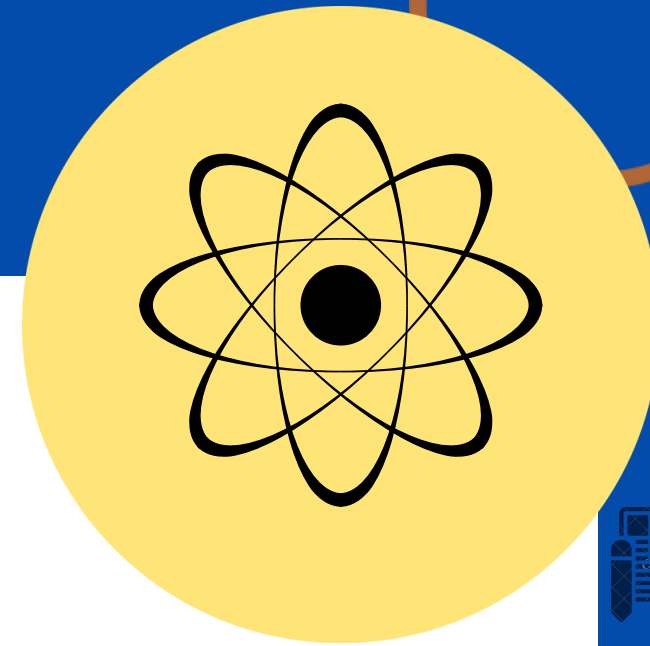
**POPULATION
CLASSIFICATIONS**

**COMMON
HOUSEHOLD
CHARACTERISTICS**



**REGIONAL-
DIVISION
SIGNIFICANCE**

**POPULATION
CLASSIFICATIONS**



DATA SOURCE ?

**COMMON
HOUSEHOLD
CHARACTERISTICS**





Independent Statistics & Analysis
U.S. Energy Information
Administration

[+ Sources & Uses](#)

[+ Topics](#)

[+ Geography](#)

[CONSUMPTION & EFFICIENCY](#)

RESIDENTIAL ENERGY CONSUMPTION SURVEY (RECS)

[OVERVIEW](#)

[DATA ▼](#)

[ANALYSIS & PROJECTIONS](#)

2015 RECS Survey Data

[2020](#) | [2015](#) | [2009](#) | [2005](#) | [2001](#) | [1997](#) | [1993](#) | [PREVIOUS](#)

[Housing characteristics](#)

[Consumption & expenditures](#)

[Microdata](#)

[Methodology](#)



Over 5,000 Observations

4	SAS Variable Name	Type	Length	Variable Description	Final Response Set
755	NGAMOUNT	Numeric	8	Concentrated energy used for domestic natural gas, 2015	Thousand Btu per gallon used for natural gas (warmed)
756	PERIODNG	Numeric	8	Number of days covered by Energy Supplier Survey natural gas billing data and used to calculate annual consumption and expenditures	1 330 days or more 2 At least 146, but less than 330 days 3 At least 60, but less than 146 days 4 Less than 60 days 5 No billing data collected or billing data not used -2 Not applicable, no natural gas consumption
757	ZNGAMOUNT	Numeric	8	Imputation flag for total natural gas usage	1 Imputed (No billing data collected or billing data not used) 0 Not imputed -2 Not applicable, no natural gas consumption
758	FOXBTU	Numeric	8	Fuel oil/kerosene to Btu conversion factor, 2015	137.45 Thousand Btu per gallons of #2 fuel oil 135 Thousand Btu per gallons of kerosene 135.48, 136.79, Thousand Btu per gallons of other fuels 137.37
759	PERIODFO	Numeric	8	Number of days covered by Energy Supplier Survey fuel oil/kerosene billing data and used to calculate annual consumption and expenditures	1 365 days (complete) 2 Less than 365 days (incomplete) 5 No delivery data collected -2 Not applicable, no fuel oil consumption
760	ZFOAMOUNT	Numeric	8	Imputation flag for total fuel oil/kerosene usage	1 Imputed (No billing data collected or billing data not used) 0 Not imputed -2 Not applicable, no fuel oil consumption
761	LPXBTU	Numeric	8	Propane to Btu conversion factor, 2015	91.33 Thousand Btu per gallons of propane
762	PERIODLP	Numeric	8	Number of days covered by Energy Supplier Survey propane billing data and used to calculate annual consumption and expenditures	1 365 days (complete) 2 Less than 365 days (incomplete) 5 No delivery data collected -2 Not applicable, no propane consumption
763	ZLPAMOUNT	Numeric	8	Imputation flag for total propane usage	1 Imputed (No billing data collected or billing data not used) 0 Not imputed -2 Not applicable, no propane consumption

Over 700 Independent Variables

Methodologies

Machine learning for the company to you is a game of Spot the Difference.



Spot the Difference

Program



KNN

Program



KNN

Program



Random Forest

KNN

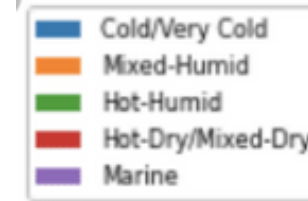
Program

Linear Regression

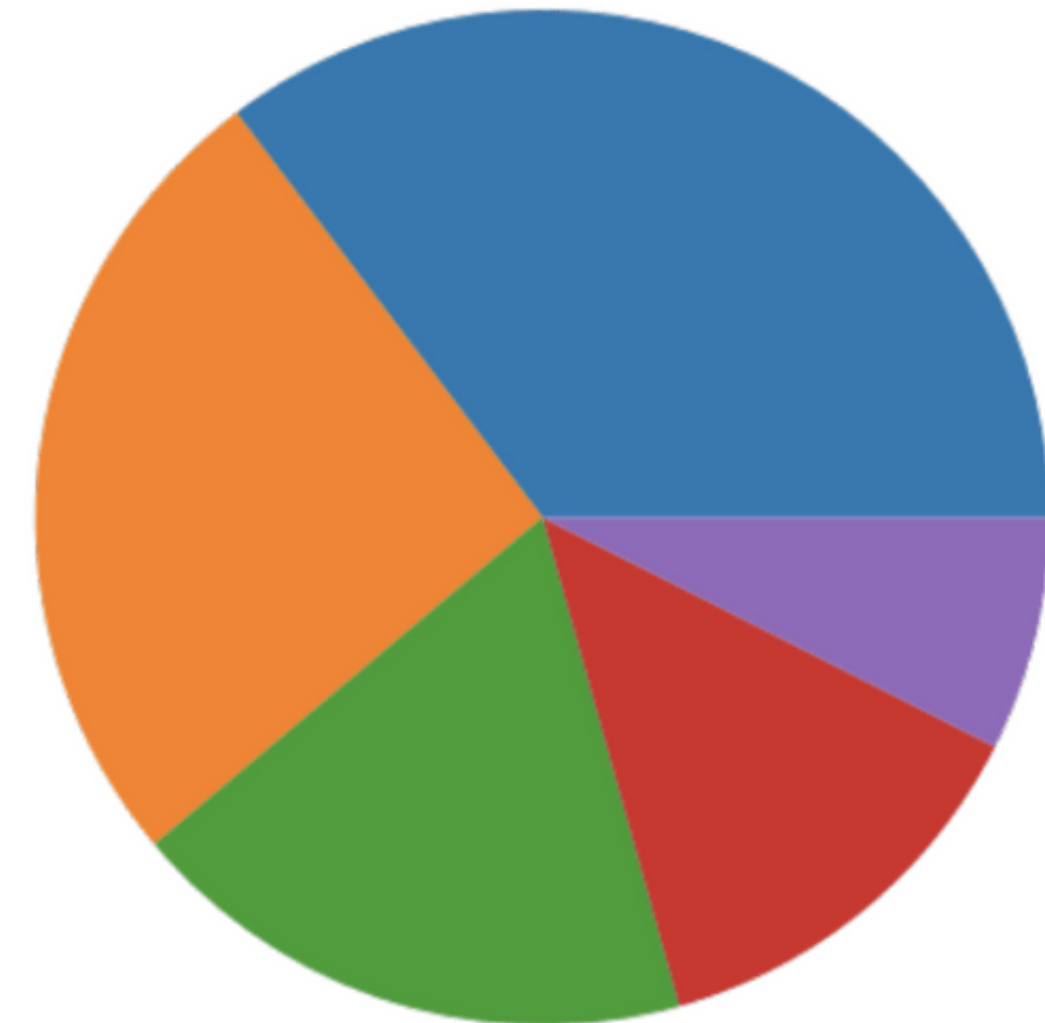
Random Forest



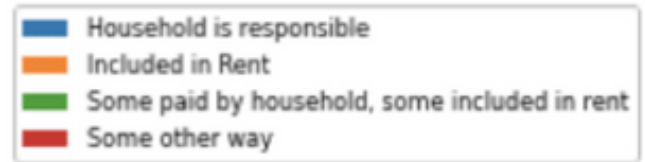
First Glance



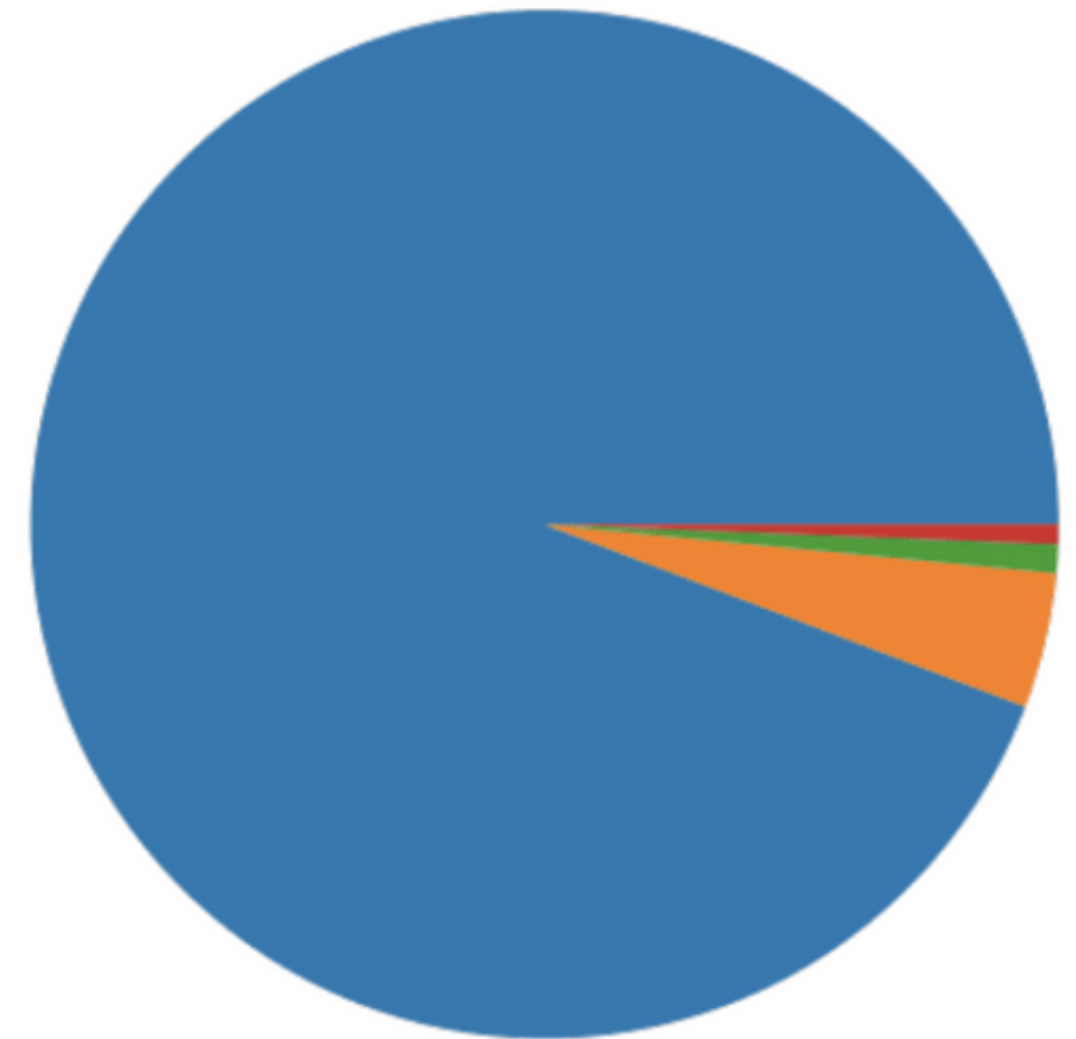
United States Climate Region Proportions



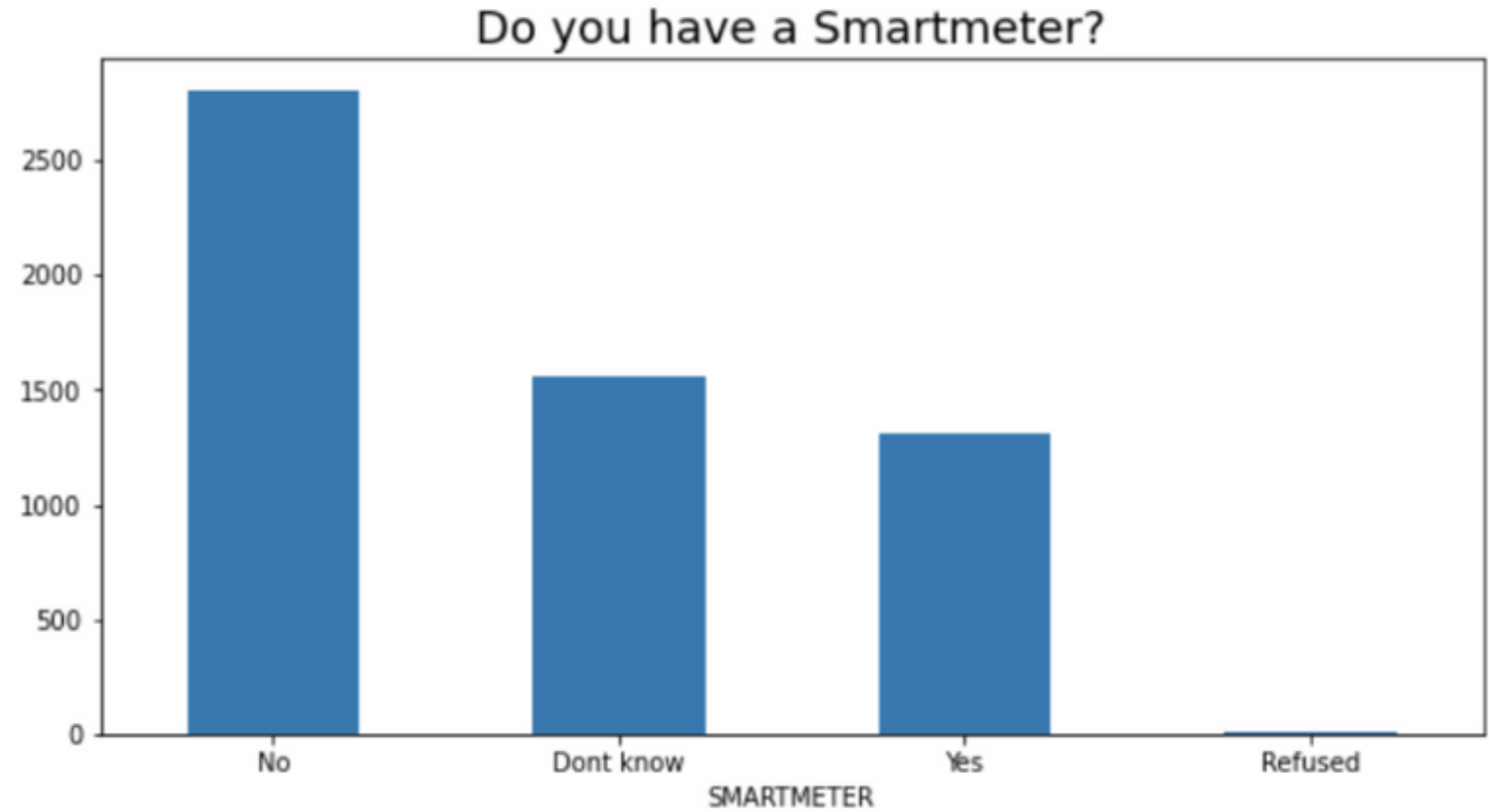
First Glance



Payment Type of Electricity Consumption

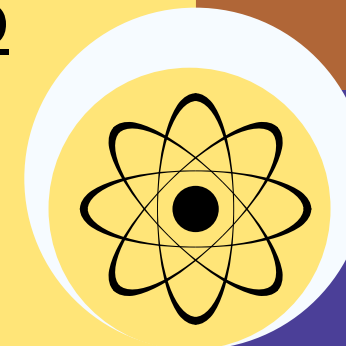


First Glance



Data Cleaning

- **Removed Imputation flags and other redundant columns.**
- **Filtered out more columns based on domain knowledge.**
- **Excluded object columns**
- **Performed Variance Inflation Factor (VIF) twice to remove multicollinearity.**
- **Created dummy variables for several categorical columns + remove outliers.**

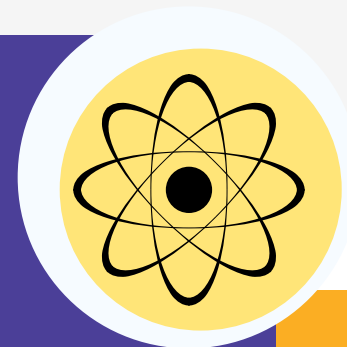


**759 variables
to 50 variables**

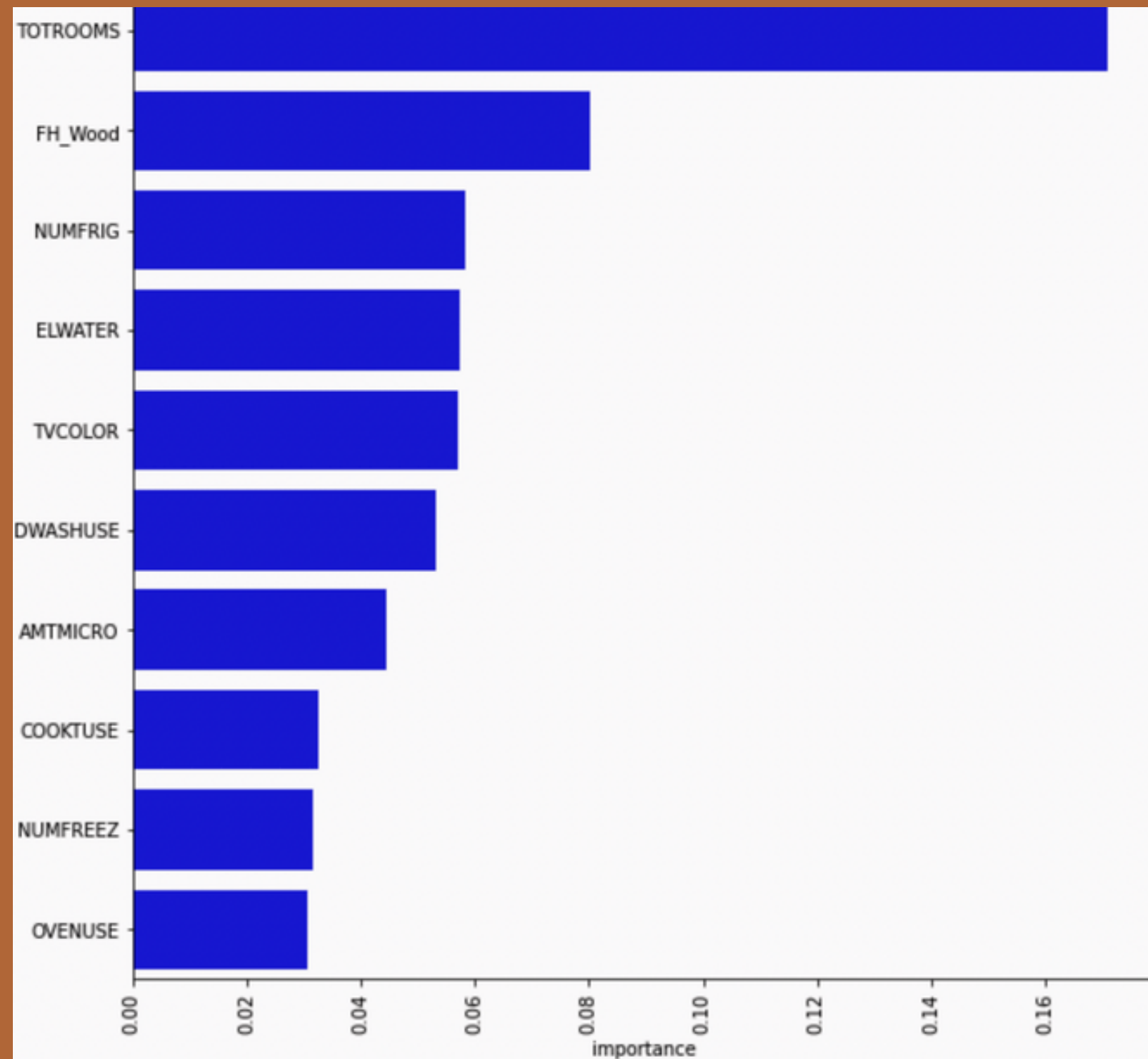
Model Accuracy

	Model	Accuracy
0	Random Forest Regressor	0.797544
1	KNN Regressor	0.713176
2	Decision Regressor	0.551552

Why?



Feature Importance



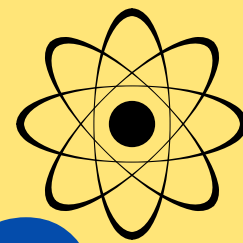
Definitions

- **TOTROOMS**- total number of rooms in housing unit , excluding bathrooms
- **FH_WOOD** - Wood used as fuel for main space heating.
- **NUMFRIG**- number of refrigerators used
- **ELWATER**- electricity used in water heating.
- **TVCOLOR**- number of televisions used
- **DWASHUSE**- frequency of dishwasher use
- **AMTMICRO**-frequency of microwave use
- **COOKTUSE**- frequency of use of cooktop part of the stove
- **NUMFREEZ**- number of separate freezers used.
- **OVENUSE**- frequency of use of oven part of stove

The Plan

Key

Takeaways *The Plan*



When do you notice
something is wrong?

When do you notice
something is wrong?

**It goes beyond predicting a bill but
using data to make a smart home.**

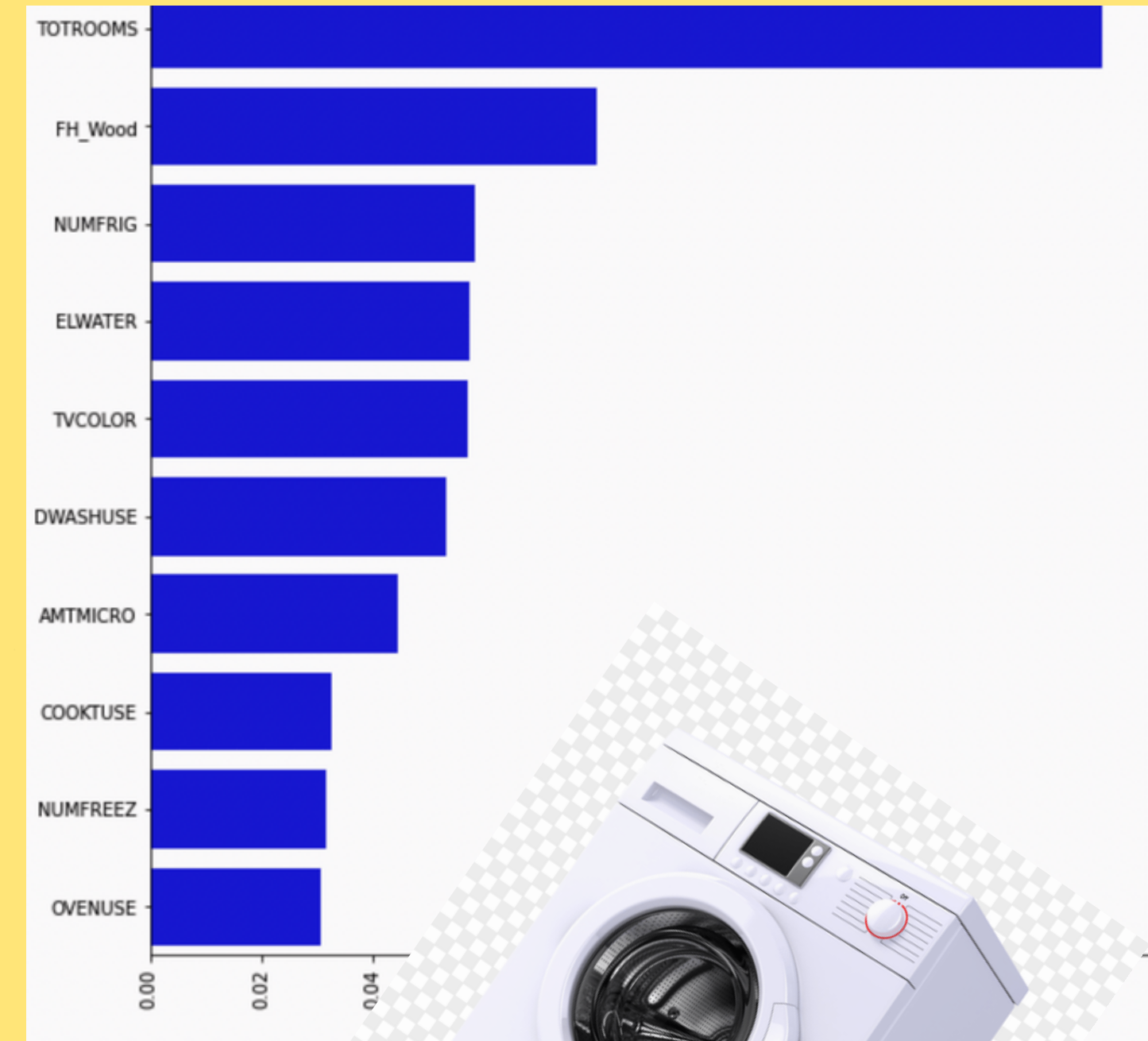


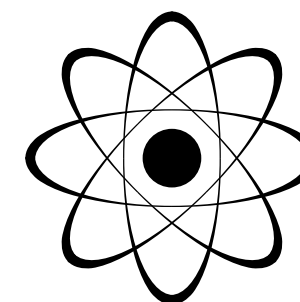
Spot the Difference

Nucleus

4 ways our model focus on you

- **Electronic Devices**
- **Smart meter +
(AutoHomeTech)**
- **Self-Education**
- **Save Money**





IMPROVEMENTS



DATA CLEANING

DUMMY VARIABLES FOR ALL
CATEGORICAL VALUES



MODELS

COMBINING DIFFERENT MODELS
TOGETHER



KNOWLEDGE

WE ONLY KNOW WHAT WE DO NOT
KNOW.

Questions?