|  |  |
| --- | --- |
| **Name** | **MAHESWARAN.S** |
| **REG .NO** | **420121104031** |
| **Department** | **Computer Science** |
| **Year** | **III** |
| **College Name** | **AKTMCET** |
| **Group** | **IBM-GROUP-5** |

Measure Energy Consumption

**PHASE-2**

**Innovation**

Innovation can definitely play a role in measuring and reducing energy consumption. Various technological advancements and creative solutions have been developed to make processes more energy-efficient and sustainable. For instance, smart meters and IoT devices can be used to monitor energy usage in real-time, providing valuable data for optimizing consumption.

Additionally, innovations in renewable energy sources, energy storage technologies, and smart grids contribute to a more sustainable and efficient energy ecosystem. The development of energy-efficient appliances, buildings, and transportation systems also stems from innovative thinking.

It's fascinating to see how creativity and technological breakthroughs can lead to more responsible and sustainable use of resources. If you have any specific innovations in mind or want more information

**PROGRAM CODE**

import pandas as pd

import requests

import json

def get\_census\_housing\_content(acs\_year):

    url = "https://api.census.gov/data/" + acs\_year + "/acs5?get=B25001\_001E,B25001\_001M,B25002\_002E,B25002\_002M,B25002\_003E,B25002\_003M,B25040\_001E,B25040\_001M,B25040\_002E,B25040\_002M,B25040\_003E,B25040\_003M,B25040\_004E,B25040\_004M,B25040\_005E,B25040\_005M,B25040\_006E,B25040\_006M,B25040\_007E,B25040\_007M,B25040\_008E,B25040\_008M,B25040\_009E,B25040\_009M,B25040\_010E,B25040\_010M&for=zip%20code%20tabulation%20area:\*&key=<API KEY HERE>"

    census\_response = requests.get(url)

    return census\_response.text

postgres\_engine = create\_engine("postgresql://<DB Credentials here>")

acs\_year = '2015'

census\_json = json.loads(get\_census\_housing\_content(acs\_year))

df = pd.DataFrame.from\_dict(census\_json,orient='columns')

df.insert(0,'year',acs\_year)

df = df.iloc[1:]

df.columns=[

'year','total\_units','total\_units\_me','occupied\_units','occupied\_units\_me',

'vacant\_units','vacant\_units\_me','total\_fuel','total\_fuel\_me','total\_fuel\_utility\_gas',

'total\_fuel\_utility\_gas\_me','total\_fuel\_bottled\_gas','total\_fuel\_bottled\_gas\_me',

'total\_fuel\_electricity','total\_fuel\_electricity\_me','total\_fuel\_oil','total\_fuel\_oil\_me',

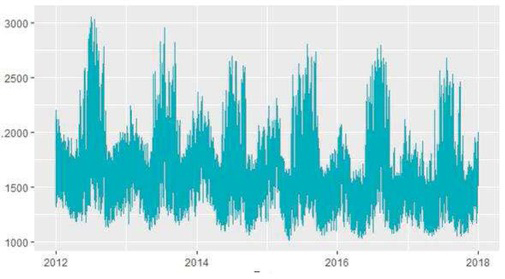
'total\_fuel\_coal','total\_fuel\_coal\_me','total\_fuel\_wood','total\_fuel\_wood\_me','total\_fuel\_solar',

'total\_fuel\_solar\_me','total\_fuel\_other','total\_fuel\_other\_me','total\_fuel\_none','total\_fuel\_none\_me','zip\_code']

df.to\_sql(name='qsdb\_census\_housing\_characteristics', con=postgres\_engine, if\_exists = 'append', index=False)

postgres\_engine.dispose();

print('----------------Census Data Processing Complete----------------')'



**Module 3** : Data Wrangling Techniques

Ah, data wrangling, the art of cleaning and organizing raw data into a more usable form! There are various techniques to master this craft. Here are a few:

* **Handling Missing Data**:
* **Imputation**: Fill in missing values with a sensible estimate.
* **Dropping**: Remove rows or columns with missing data.
* **Data Transformation:**
* **Normalization/Scaling**: Standardize numerical values to a common scale.
* **Encoding**: Convert categorical data into numerical form for analysis.
* **Binning**: Group numerical data into bins for better analysis.
* **Dealing with Duplicates**:
* **Identifying and Removing**: Detect and eliminate identical rows.
* **Text Data Cleaning:**
* **Tokenization**: Break down text into individual words or phrases.
* **Removing Stopwords**: Eliminate common words that don't carry much meaning.
* **Stemming/Lemmatization**: Reduce words to their base or root form.
* **Handling Outliers**:
* **Detection**: Identify and examine data points significantly different from the rest.
* **Treatment**: Decide whether to remove, transform, or keep outliers based on the context.
* **Data Aggregation**:
* **Grouping**: Group data by certain criteria for analysis.
* **Summarization:** Create summary statistics for better understanding.
* **Joining and Merging:**
* **Combining Datasets:** Merge datasets based on common columns.
* **Time Series Data Techniques:**
* **Resampling:** Adjust the time frequency of the data.
* **Lagging and Leading**: Create new time-shifted features.
* **Regex (Regular Expressions):**
* **Pattern Matching:** Use regex to identify and manipulate patterns in text data.
* **Data Validation:**
* **Checking for Consistency:** Ensure data adheres to predefined rules.
* **Data Formatting:**
* **Date Formatting:** Ensure consistency in date formats.
* **Numeric Formatting:** Make sure numerical data is in the right format.

**Module 4:** Introduction to Neural Networks

Neural networks are like the cool kids of machine learning. Imagine them as a bunch of interconnected nodes, or neurons, inspired by the human brain.

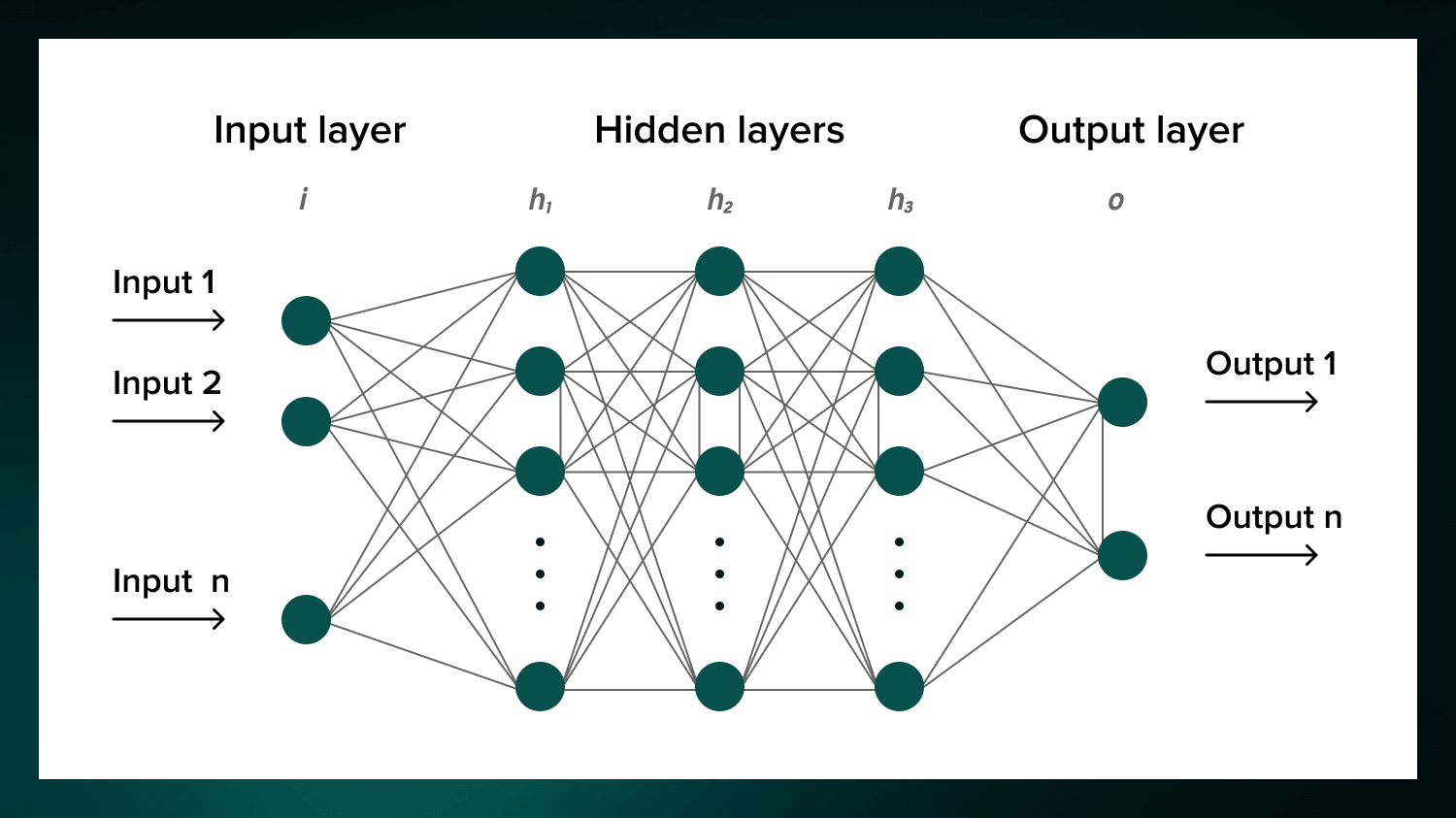
Each connection has a weight, and these weights are adjusted during training to make the network learn.

There are input layers where data goes in, output layers where results come out, and hidden layers in between doing the heavy lifting.

The magic happens when the network learns to map inputs to outputs by adjusting those weights. It's like a brain, but for computers.

Deep learning takes it up a notch with deep neural networks—more layers, more complexity.

They've been rocking the tech scene, doing everything from image recognition to natural language processing.



**Weighty matters**

Neural nets are a means of doing machine learning, in which a computer learns to perform some task by analyzing training examples. Usually, the examples have been hand-labeled in advance. An object recognition system, for instance, might be fed thousands of labeled images of cars, houses, coffee cups, and so on, and it would find visual patterns in the images that consistently correlate with particular labels.

**Minds and machines**

The neural nets described by McCullough and Pitts in 1944 had thresholds and weights, but they weren’t arranged into layers, and the researchers didn’t specify any training mechanism. What McCullough and Pitts showed was that a neural net could, in principle, compute any function that a digital computer could. The result was more neuroscience than computer science: The point was to suggest that the human brain could be thought of as a computing device.