

NumPy Project

Using NumPy to explore and analyze the data present in the publicly available file from the U.S. Energy Information Administration which contains information on electricity generation in the USA from a range of sources. The data file is a CSV file, so use the csv module to import the data into a list which should be converted to an appropriate numpy array(as shown below)

<https://drive.google.com/file/d/1Cz0Yvx4yES8ydhLaBSnclIPxW2yKHwWH/view?usp=sharing>

Link to the dataset:

Below is the code to read the above file and convert into a numpy array.
(Note to replace the path of the file in cell 4 (*with open*) according to your setup.

```
In [1]: import csv

In [2]: import numpy as np

In [4]: data = []

with open(r"D:\Downloads\MER_T07_02A-2020-02-03.csv", 'r') as csvfile:
    file_reader = csv.reader(csvfile, delimiter=',')
    for row in file_reader:
        data.append(row)

data = np.array(data) #convert the List of Lists to a NumPy array
```

The following tasks need to be performed on the above data:

1. Explore the important attributes like dimension, shape, data type etc, of the array formed above.
2. Print the data contained in the first 10 rows of the 4th column.
3. Which row serves as the headers/titles for all the columns.
4. Print the data contained in column 2 and 3 from row 1 till row 20
5. Print the data present in only the first three and the last three rows of all the columns in a single output.
6. Sort the data on the basis of net amount of electricity generated irrespective of the source.
7. Find the total amount of electricity generated using coal and nuclear between 1949-1990. (In this dataset, rows containing monthly data express date in the format 'YYYYMM'. Rows containing annual data express the date in the format 'YYYY13'.)
8. Print all the unique sources of Energy generation present in the dataset.
9. Print all the details(annual) where the energy source is Wind Energy. Use the concept of masking to filter the data.
10. Print the Total Energy generated in the USA till date.

11. Print the average annual energy generated from wind in the USA and also the standard deviation present in the energy generation.
12. What and when was the maximum annual energy generated?
13. Find from the above data if the energy production in the USA has increased in the last 10 years.
14. What is the trend in the energy generated from wind over the years? Which source of energy has been the largest contributor to the annual electricity production over the years.
15. Which renewable source of energy is expected to meet the major demand of energy in the coming years. Can we predict an estimate for the same.
16. Compute the contribution of wind, solar, and their combined contribution compared to the total energy generation in the USA, and answer whether the national grid is fundamentally shifting toward wind energy? Use numpy and appropriate metrics to answer the question.