**To Analyse and Predict the Rise of Energy Consumption and Predict Additional Energy Generation, in Future for NTPC, Ltd**

**Biswaroop Nath**

*A project report for Winter Internship Dec 2022-Jan 2023 at NTPC*

**Energy Analysis Requirement**

Electricity being used as a commodity in houses have helped people make use of it in innovative ways. It has made the lives of people easy. Hence an analysis of energy is always important.

Here NTPC has provided a major role of power production in India since 1975.

**Power Generation Requirement Analysis for NTPC**

NTPC always has been considered above average in terms of power production in India. It accounts for 25% approximately of India's total power generation. This study will give a comprehensive analysis of how much more power generation does NTPC requires to keep up this performance in the future as the demand for power supply rises due to various factors which have also been analyzed here.

**Contents**

1. Energy Analysis Requirement-------------------------------------1
2. Power Generation Requirement Analysis for NTPC--------------1
3. Data Analysis------------------------------------------------------3
4. Basic Statistical Information--------------------------------------5
5. Co-relation Analysis-----------------------------------------------9
6. Further Plots for analysis----------------------------------------10
7. Comparison of Different Models for prediction----------------12
8. Prediction plot for each factors using Linear Regression------13
9. Summary of predictions-----------------------------------------15
10. Projection of NTPC’s Requirements-----------------------------16
11. Recommendations-----------------------------------------------20
12. Conclusion-------------------------------------------------------20
13. Bibliography-----------------------------------------------------20

**Data Analysis** (Exploratory Data Analysis)

To predict the rise in demand of energy in India we should take a look and analyze various possible factors , that might be responsible. For this we need to track India’s history on multiple areas , and find and correlate if they are responsible for rise in energy demand.

To make this study simple , I have collected data from the year 1971 to 2014 , with factors :

* *Population Growth*: Might be responsible for increase in demand , because more the people the more the energy requirements.
* *Electricity Access (in percentage of population)*: Electrical access is increasing over the years as our country is developing , hence it might be good factor to calculate and analyze.
* *per Capita Electricity consumption(pCEC in KWH)*: Will denote how much energy per person uses on an average.
* *GDP:* Denotes the general development of our country.

These factors might be responsible for the rise in demand of power in India so let us analyze them.

Here is the quick look at the data.

Note: All the data-sets that have been utilized are taken from https://data.worldbank.org/country/IN and <https://ourworldindata.org>

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year** | **Energy Consumption(TWh)** | **Population Growth** | **GDP** | **Per-Capita Power Consumption** | **Electricity Access** |
| 1971 | 794.3863 | 2.257900247 | 67350988021 | 97.77271821 | 28.9 |
| 1972 | 835.8586 | 2.301189197 | 71463193831 | 100.3050738 | 29.9 |
| 1973 | 853.3237 | 2.327407855 | 85515269585 | 100.6321541 | 30.9 |
| 1974 | 903.9876 | 2.331906417 | 99525899116 | 103.9926579 | 31.9 |
| 1975 | 972.0112 | 2.321760467 | 98472796457 | 114.5589276 | 32.9 |
| 1976 | 1019.4541 | 2.304663743 | 1.03x1011 | 124.1236916 | 33.9 |
| 1977 | 1078.32 | 2.291299608 | 1.21x1011 | 126.256734 | 34.9 |
| 1978 | 1120.9728 | 2.286780592 | 1.37x1011 | 135.9176407 | 35.9 |
| 1979 | 1182.959 | 2.29489353 | 1.53x1011 | 135.7417737 | 36.9 |
| 1980 | 1218.599 | 2.309521943 | 1.86x1011 | 141.7077015 | 37.9 |
| 1981 | 1349.2845 | 2.323758727 | 1.93x1011 | 151.8790588 | 38.9 |
| 1982 | 1336.0491 | 2.328678839 | 2.01x1011 | 158.0999117 | 39.9 |
| 1983 | 1405.0502 | 2.320388368 | 2.18x1011 | 165.687486 | 40.9 |
| 1984 | 1500.9186 | 2.295805502 | 2.12x1011 | 183.3018723 | 41.9 |
| 1985 | 1584.2627 | 2.259853172 | 2.33x1011 | 193.5348535 | 42.9 |
| 1986 | 1695.5906 | 2.220963311 | 2.49x1011 | 207.976493 | 43.9 |
| 1987 | 1802.8536 | 2.183882902 | 2.79x1011 | 220.220743 | 44.9 |
| 1988 | 1962.3015 | 2.146759799 | 2.97x1011 | 240.0256187 | 45.9 |
| 1989 | 2143.4136 | 2.110865104 | 2.96x1011 | 257.0397371 | 46.9 |
| 1990 | 2300.7012 | 2.076089204 | 3.21x1011 | 272.0634834 | 47.9 |
| 1991 | 2427.663 | 2.039728736 | 2.70x1011 | 290.8995799 | 48.9 |
| 1992 | 2552.7722 | 2.003178624 | 2.88x1011 | 304.4318305 | 49.9 |
| 1993 | 2616.248 | 1.970634987 | 2.79x1011 | 320.5507448 | 50.9 |
| 1994 | 2757.1213 | 1.943243922 | 3.27x1011 | 341.2313623 | 49.81130981 |
| 1995 | 2963.118 | 1.918940659 | 3.60x1011 | 358.7622129 | 51.40877533 |
| 1996 | 3081.6313 | 1.895219577 | 3.93x1011 | 359.8162707 | 53.00352097 |
| 1997 | 3252.01 | 1.869172003 | 4.16x1011 | 375.4850529 | 54.59486389 |
| 1998 | 3449.4966 | 1.839658764 | 4.21x1011 | 385.8639858 | 56.18213272 |
| 1999 | 3549.4888 | 1.805559737 | 4.59x1011 | 392.039693 | 60.1 |
| 2000 | 3728.5144 | 1.76812551 | 4.68x1011 | 393.6462478 | 58.72147369 |
| 2001 | 3743.8525 | 1.72876857 | 4.85x1011 | 393.8101981 | 55.79999924 |
| 2002 | 3865.7512 | 1.689561661 | 5.15x1011 | 410.6447839 | 62.29999924 |
| 2003 | 4004.0781 | 1.651491269 | 6.08x1011 | 430.4831633 | 64.04748535 |
| 2004 | 4347.582 | 1.615308295 | 7.09x1011 | 451.6115461 | 64.40000153 |
| 2005 | 4603.6104 | 1.579709143 | 8.20x1011 | 468.025754 | 67.5798111 |
| 2006 | 4849.72 | 1.545696439 | 9.40x1011 | 509.2140548 | 67.90000153 |
| 2007 | 5263.1113 | 1.509221986 | 1.22x1012 | 541.7383952 | 71.11986542 |
| 2008 | 5570.697 | 1.464889915 | 1.20x1012 | 561.2475814 | 72.89938354 |
| 2009 | 6003.236 | 1.410582714 | 1.34x1012 | 598.4982419 | 75 |
| 2010 | 6269.459 | 1.350338314 | 1.68x1012 | 640.3946068 | 76.30000305 |
| 2011 | 6650.7104 | 1.288512962 | 1.82x1012 | 696.8426815 | 67.59999847 |
| 2012 | 7048.088 | 1.231484894 | 1.83x1012 | 723.2369166 | 79.90000153 |
| 2013 | 7294.344 | 1.182904215 | 1.86x1012 | 764.2011341 | 81.99932861 |
| 2014 | 7799.773 | 1.145673402 | 2.04x1012 | 804.5163493 | 83.87249756 |

**Table 1: For data analysis on various factors.**

**Basic Statistical Information**

Let us calculate some basic fundamental statistical information about our dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Attrib-ute** | **Year** | **Energy Consumption(TWh)** | **Population Growth** | **GDP** | **Per-Capita Power Consumption** | **Electricity Access** |
| count | 44 | 44 | 44 | 4.40x101 | 44 | 44 |
| mean | 1992.5 | 3062.553964 | 1.925954 | 5.54x1011 | 335.182516 | 52.096374 |
| std | 12.845233 | 2002.485011 | 0.374892 | 5.59x1011 | 198.598853 | 15.472843 |
| min | 1971 | 794.3863 | 1.145673 | 6.74x1010 | 97.772718 | 28.9 |
| 25% | 1981.75 | 1345.97565 | 1.642446 | 1.99x1011 | 156.544698 | 39.65 |
| 50% | 1992.5 | 2584.5101 | 1.986907 | 3.09x1011 | 312.491288 | 49.855655 |
| 75% | 2003.25 | 4089.954075 | 2.292198 | 6.33x1011 | 435.765259 | 64.135614 |
| max | 2014 | 7799.773 | 2.331906 | 2.04x1012 | 804.516349 | 83.872498 |

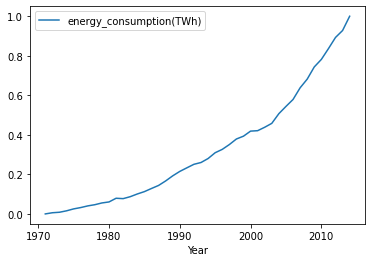
**Table 2: Statistical information.**EC(energy consumption)

As we can clearly see the data values are too high for some data sets and the scaling also varies , hence it is wiser to use *min-max normalization* to normalize the data so that certain factors with larger values are automatically not considered as important variables , hence making the data analysis much fair and accurate.

There is also another benefit of Min-Max Normalization, the data gets evenly distributed, without loosing the nature of the curve of the attribute.

Let us look at the plot of each attribute to understand and visualize the general trend of it.

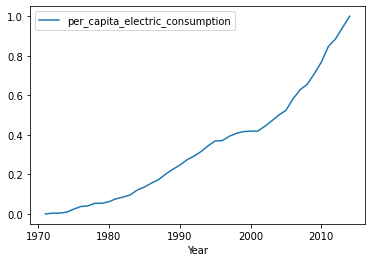
* Absolute Power Consumption



**Fig.1:Plot for Absolute power consumption (in BU)**

We can see it has an increasing trend over the years.

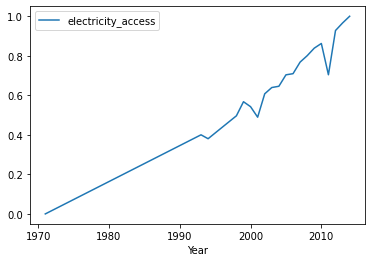
* Per-Capita Energy Consumption



**Fig.2.Plot for Per-Capita Energy Consumption**

Per capita-energy consumption is also increasing in nature , as expected.

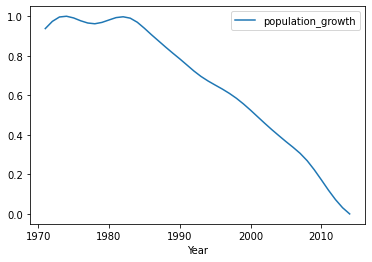
* Electricity-Access



**Fig.3.Plot for Electricity Access**

Electricity access has some dips in graphs but there is a general increasing trend.

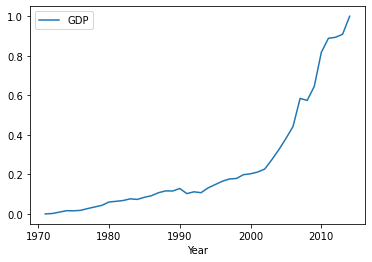
* Population Growth



**Fig.4.Plot for Population Growth**

Clearly, India’s population growth has been declining since after the 1980’s.

* GDP



**Fig.5.Plot for GDP**

Clearly, the GDP is also have a steep increasing trend.

**Co-relation Analysis**

For us to understand if the factors are correlated with Energy consumption, and if they would be a good for predicting future demand of energy, we analyze, if the factors are interdependent or not. This can be done by two ways:

* Cosine Similarity

|  |  |  |
| --- | --- | --- |
| **Cosine Similarity** | **Absolute Energy consumption** | **Inference** |
| Electricity Access | 0.98 | Highly Related |
| Population Growth | 0.40 | Less Related |
| GDP | 0.97 | Highly Related |
| pCEC | 0.99 | Highly Related |

**Table 3: Cosine Similarity on various factors.**

* Pearson's Co-relation Co-efficient

|  |  |  |
| --- | --- | --- |
| **Pearson’s correlation coefficient** | **Absolute Energy consumption** | **Inference** |
| Electricity Access | 0.979 | Positively Co-related |
| Population Growth | -0.992 | Negatively Co-related |
| GDP | 0.960 | Positively Co-related |
| pCEC | 0.997 | Positively Co-related |

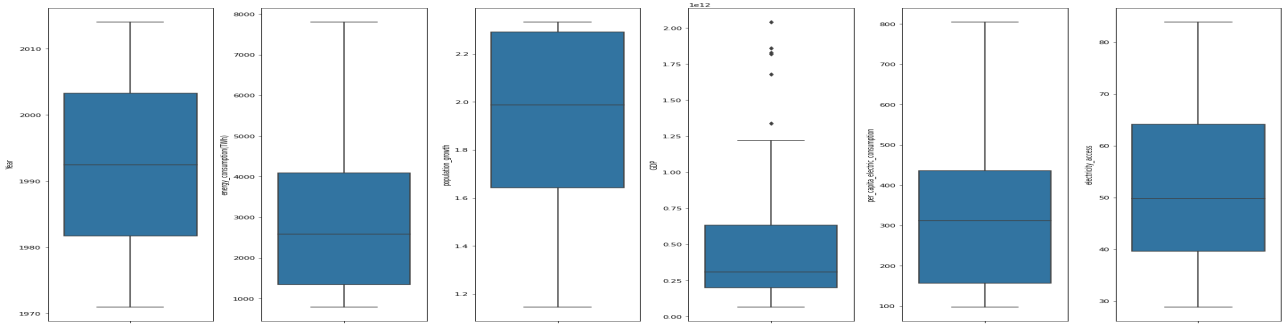
**Table 4: Pearson’s Corelation Coefficient on various factors.**

* Conclusions

After this correlation analysis, we can confidently conclude that GDP, Electrical Access , Per Capita Energy Consumption are highly related.

**Further Plots for analysis**

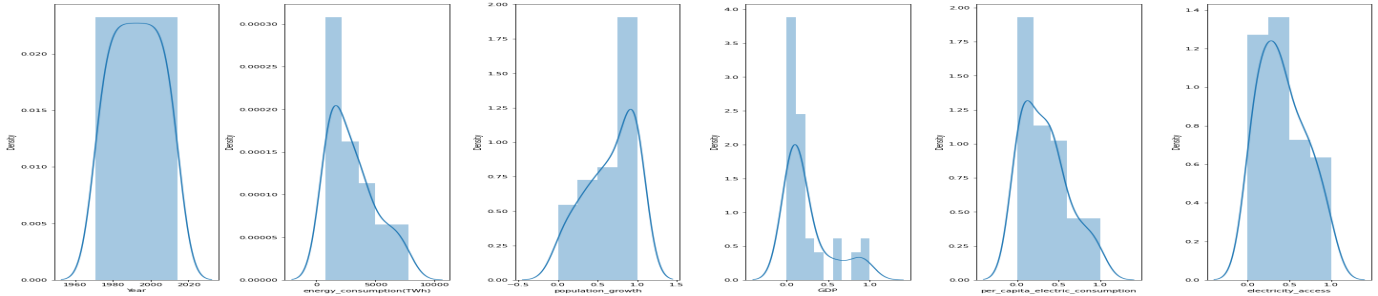
* Boxplot



**Fig.6.BoxPlot**

This analysis shows us that there aren't any significant outliers in our attributes.

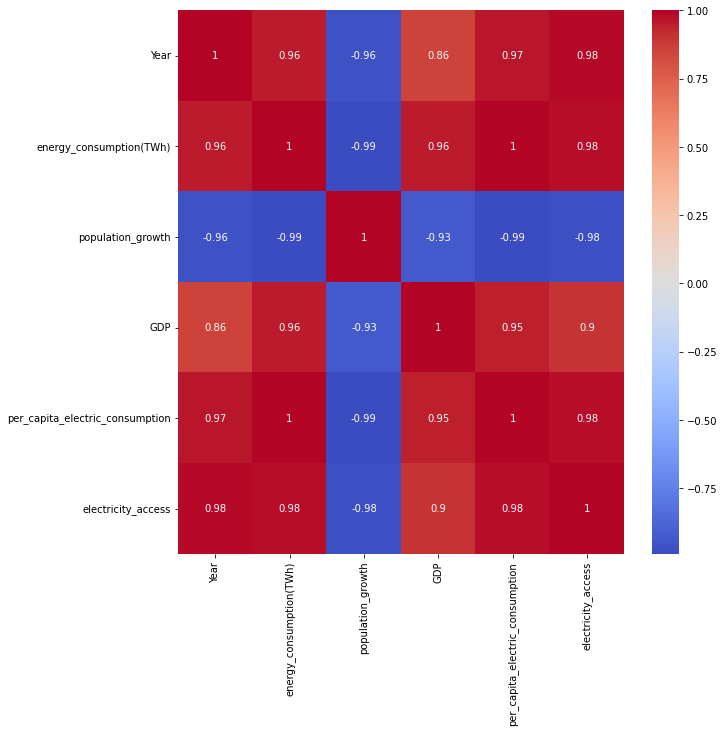
* Normalized Distplot



**Fig.7.Normalized DistPlot**

Since we have normalized the data before-hand,the data is evenly distributed.

* Heatmap/Cor-relation Matrix



**Fig.8.Correlation Matrix**

This is the correlation matrix showing correlations between all the attributes. We can clearly infer that population growth is negatively correlated.

**Comparison of Different Models for prediction**

We have to select which model to train on using our dataset. Now there are multiple models available, I have trained them on some of the most popular models on our dataset and have split the data into train and test samples to compare, which ever-one having the least error, we will select that model for final predictions.

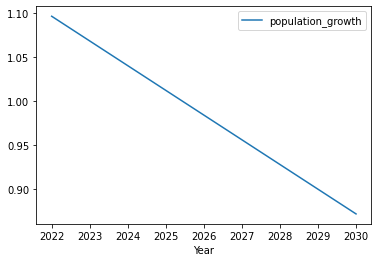
|  |  |
| --- | --- |
| **Model** | **MAE(Mean Absolute Error)** |
| Linear Regression | 222.25 |
| Decision Tree | 928.76 |
| Random Forest | 1181.54 |
| Extra Trees Regressor | 1068.71 |
| Extreme Gradient Boosting | 955.35 |

**Table 5: Comparison of various Models.**

*Conclusion: The least MAE is from Linear Regression only so we will use linear regression model.*

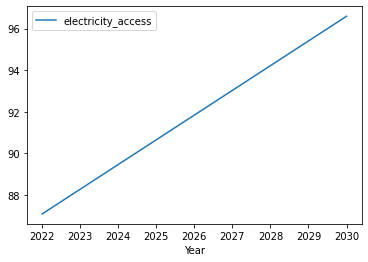
**Prediction plot for each factors using Linear Regression**

Since Linear Regression is the model we picked from our analysis, we train it to predict and provide us general trend over the upcoming years.



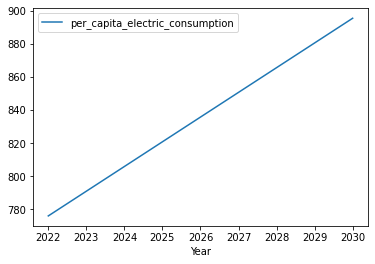
**Fig.9.Prediction plot for population growth.**

As expected the population growth is declining.



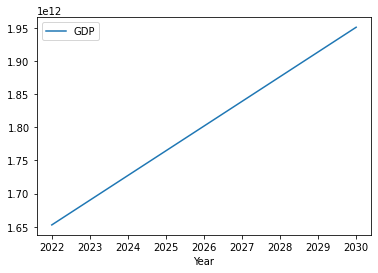
**Fig.10.Prediction plot for Electricity Access.**

Also, electricity access is increasing.



**Fig.11.Prediction plot for pCEC.**

We can clearly infer per capita energy consumption is increasing as well.



**Fig.12.Prediction plot for GDP.**

The GDP attribute is also increasing.

**Summary of predictions**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **In** | **Year** | **pG** | **GDP** | **pCEC** | **EA** |
| 0 | 2022 | 1.096489 | 1.652768x1012 | 775.880197 | 87.09768 |
| 1 | 2023 | 1.068371 | 1.690019x1012 | 790.819101 | 88.284165 |
| 2 | 2024 | 1.040254 | 1.727271x1012 | 805.758005 | 89.470649 |
| 3 | 2025 | 1.012136 | 1.764522x1012 | 820.69691 | 90.657134 |
| 4 | 2026 | 0.984019 | 1.801774x1012 | 835.635814 | 91.843619 |
| 5 | 2027 | 0.955902 | 1.839025x1012 | 850.574719 | 93.030104 |
| 6 | 2028 | 0.927784 | 1.876277x1012 | 865.513623 | 94.216589 |
| 7 | 2029 | 0.899667 | 1.913528x1012 | 880.452527 | 95.403074 |
| 8 | 2030 | 0.871549 | 1.950780x1012 | 895.391432 | 96.589559 |

**Table 6: Prediction : Values for various factors.**

*In(Index),pG(population growth),pCEC(per-capita-energy-consumption),EA(electricity access)*

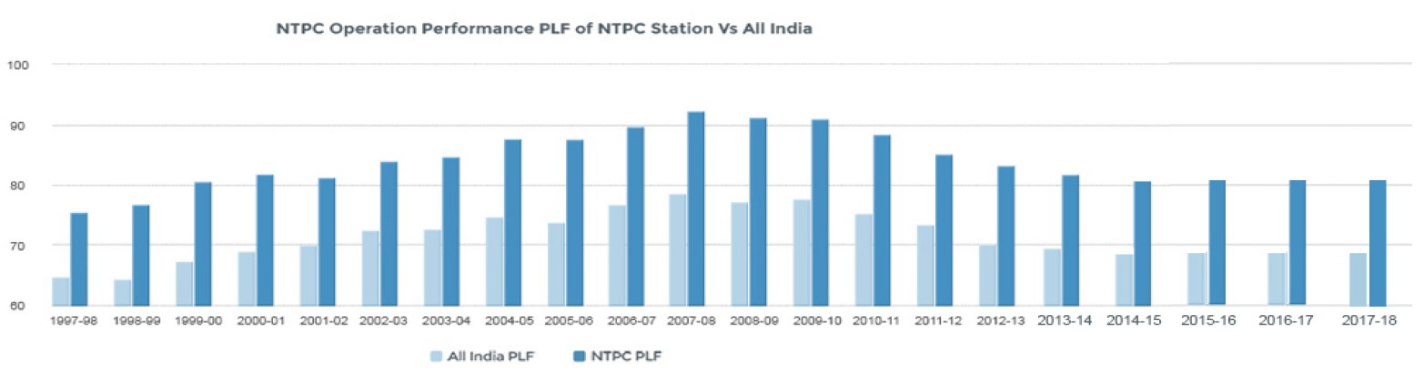
This is just the predicted values for the upcoming years for each factors, now we will input it to our model (after normalizing them) to predict energy demand in future.

**Projection of NTPC’s Requirements**

NTPC annually contributes approximately 25% of total power generated in India. To maintain a similar performance in the future the following projection has been performed to estimate how much capacity of total energy is needed to be produced by NTPC in the upcoming years.

As of 2022-2023, NTPC has generated 300 BU of electricity, we will take this as a base line in our model for predicting future additional energy requirements.

The below graph denotes the current performance of NTPC compared to India as a whole.



**Fig.13.Comparison of NTPC performance with all of India.**

*Note: The data has been taken from <https://www.ntpc.co.in/en/power-generation/performance-statistics>*

We use our linear regression model to predict the energy demand.

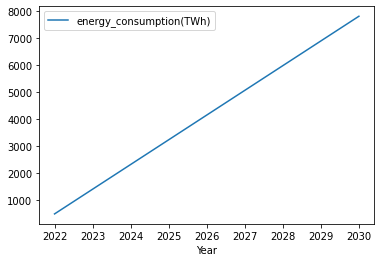
* Energy demand prediction in future

After inputting our predicted values (normalized) for various attributes, we get the below table as the predicted future demand.

|  |  |
| --- | --- |
| **Year** | **Energy Consumption(TWh)** |
| 2023 | 1395.790586 |
| 2024 | 2310.234446 |
| 2025 | 3224.678306 |
| 2026 | 4139.122165 |
| 2027 | 5053.566025 |
| 2028 | 5968.009885 |
| 2029 | 6882.453744 |
| 2030 | 7796.897604 |

**Table 7: Prediction :Future energy demand.**

Let us visualize it.



**Fig.14.Plot for future energy demand.**

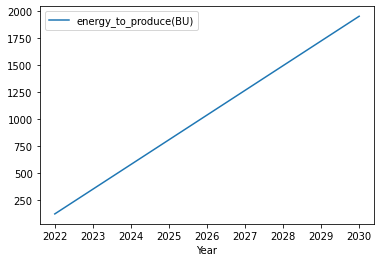
* Prediction of NTPC’s share

To maintain the current performance , NTPC needs their share of energy production as tabulated below. Since NTPC closely contributes to 1/4th of India’s power requirements , we simply find 25% of all India’s energy demand to find NTPC’s share.

|  |  |
| --- | --- |
| Year | Energy to Produce(BU) |
| 2023 | 348.947647 |
| 2024 | 577.558612 |
| 2025 | 806.169576 |
| 2026 | 1034.780541 |
| 2027 | 1263.391506 |
| 2028 | 1492.002471 |
| 2029 | 1720.613436 |
| 2030 | 1949.224401 |

**Table 8: Prediction for NTPC’s future energy share.**

Again let us visualize this in form of a graph.



**Fig.15.Plot for NTPC’s future energy share.**

* Additional future production

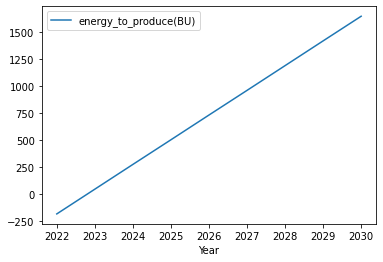
Let us calculate how much further they need to produce in future to keep up their performance.

*Note:The additional production of energy has been calculated by considering the fact that NTPC can currently produce 300 BU annually.*

|  |  |
| --- | --- |
| Year | **Additional Energy to Produce(BU)** |
| 2023 | 49 |
| 2024 | 278 |
| 2025 | 506 |
| 2026 | 735 |
| 2027 | 963 |
| 2028 | 1192 |
| 2029 | 1421 |
| 2030 | 1649 |

**Table 9: Prediction for NTPC’s additional future energy production.**

The visualization of the above table.



**Fig.16.Plot for NTPC’s additional future energy production.**

**Recommendations**

To maintain NTPC’s current performance, NTPC should consider producing *atleast* 1600 BU more by the end of 2030 (in comparison to current production) to keep up with its current standards, considering several factors that will further lead to increase in demand of power in future.

**Conclusion**

India is a developing country with 2nd largest population in the world and with the 5th largest GDP(currently 2022) , has huge demand of power for industrialization and development of the country. Hence, NTPC should strive to provide as much production of power possible for the development of our mother land, India.

**Bibliography**

1. <https://ourworldindata.org>
2. <https://www.ntpc.co.in/en/power-generation/performance-statistics>
3. <https://data.worldbank.org/country/IN>
4. <https://medium.com/analytics-vidhya/understanding-the-linear-regression-808c1f6941c0>
5. <https://www.geeksforgeeks.org/pearson-correlation-coefficient/>
6. <https://www.geeksforgeeks.org/cosine-similarity/>
7. <https://www.geeksforgeeks.org/data-normalization-in-data-mining/>