IN-DEPTH ANALYSIS OF ENERGY CONSUMPTION ACROSS VARIOUS REGIONS USING PANDAS DATA ANALYSIS TOOLS

Meliboev Azizjon Ikromjon o‘g‘li,
Department of Digital Technology and Mathematics,
Kokand University

Abstract: This paper presents a comprehensive analysis of energy consumption patterns across various regions utilizing the Pandas data analysis library in Python. The primary objective is to identify underlying trends, regional disparities, and key factors influencing energy usage by examining a rich dataset encompassing variables such as population density, industrial activities, climatic conditions, and socioeconomic indicators. Leveraging the robust data manipulation and analytical capabilities of Pandas, we efficiently processed large-scale datasets, performed intricate calculations, and generated insightful visualizations. Our findings reveal significant variations in energy consumption correlated with economic development levels, climate zones, and industrialization rates among the regions studied. These insights offer valuable implications for policymakers and energy sector stakeholders aiming to formulate targeted strategies for enhancing energy efficiency and promoting sustainable practices. The study underscores the efficacy of using Pandas as a powerful tool for conducting large-scale data analyses in energy consumption research.

Key words: Data analysis, Pandas tool, EDA, demographic analysis.

Introduction:

Demographic analysis is a crucial tool for understanding the composition, distribution, and trends of

Energy consumption is a pivotal factor in the socioeconomic development of nations and has significant implications for environmental sustainability. As global energy demands continue to escalate due to population growth, industrialization, and technological advancements, understanding the patterns and determinants of energy usage becomes increasingly critical. An in-depth analysis of energy consumption across various regions can unveil disparities, identify inefficiencies, and highlight opportunities for implementing more sustainable energy practices.

Historically, energy consumption studies have employed a variety of analytical methods ranging from econometric modeling to statistical analysis. While these methods have provided valuable insights, they often require specialized software and can be limited in handling large, complex datasets. The advent of big data and the increasing availability of open-source tools have transformed the landscape of data analysis. In particular, the Python-based Pandas library has emerged as a powerful tool for data manipulation and analysis, offering robust functionalities for handling large datasets efficiently.

Pandas facilitates the importing, cleaning, and preprocessing of data, enabling researchers to perform comprehensive analyses without the steep learning curve associated with more traditional statistical software. Its integration with other Python libraries also allows for advanced data visualization and machine learning applications, making it an ideal choice for multidisciplinary studies involving large amounts of data.

This paper leverages the capabilities of Pandas to conduct a comprehensive analysis of energy consumption patterns across various regions. By integrating datasets that include variables such as population density, industrial output, climatic conditions, and socioeconomic indicators, we aim to uncover the underlying factors that influence energy usage.

To demonstrate the effectiveness of Pandas in large-scale data analysis, we aim to showcase how Pandas can be utilized to efficiently process and analyze complex datasets, highlighting its utility in energy consumption research.

Related works

Demographic analysis has long been a subject of interest for social scientists, economists, and pas it Numerous studies have been conducted on the energy consumption of Uzbekistan, with a focus on various area aspects such as Industy, Transportation. Traditionally, energy consumption studies in Uzbekistan have been conducted using census data and statistical techniques. However, with the increasing availability of datasets and advanced data analysis tools, such as Python’s Pandas library [1], the field of energy consumption analysis has seen significant advancements in recent years.

The analysis of energy consumption patterns has been a significant area of research due to its implications for economic development, environmental sustainability, and policy-making. Traditional studies have often employed statistical and econometric models to explore the relationship between energy consumption and various factors such as economic growth, population dynamics, industrial activities, and climatic conditions. These studies have provided valuable insights into how energy usage correlates with socioeconomic indicators and have informed strategies for energy efficiency and conservation. Several works have utilized data science tools to

The Tunisian Company of Electricity and Gas (STEG) is a public entity operating independently of government administration, tasked with supplying electricity and gas throughout Tunisia in the work [2]. The company experienced significant losses amounting to 200 million Tunisian Dinars due to fraudulent tampering with meters by consumers.

Utility companies play a vital role in developed societies by providing essential services such as electricity, gas, and other forms of energy to homes, businesses, and industries in the [3].

The infrastructure required to deliver these services includes extensive networks of pipelines and electrical lines that transport energy to millions of consumers. This infrastructure needs to be safeguarded and maintained against various risks, including both technical and non-technical losses, such as fraud.

One significant challenge for these companies is the discrepancy between the energy delivered and the energy billed. This issue, known as energy losses, can result from fraud or technical factors like faulty meters, damaged pipes, or other operational problems.

In doing so, the paper builds upon the existing body of knowledge by combining methodological advancements in data analysis with substantive insights into energy consumption. It underscores the importance of using versatile data analysis tools to handle the complexities of large datasets and demonstrates how such tools can enhance the robustness and scope of energy studies. This contribution is particularly relevant in the context of growing global energy demands and the pressing need for sustainable energy solutions informed by comprehensive data analysis.

Method

The data is essential resource for the analising and get varios insight. we used statistic data which

The collected data by Statistics Agency under the President of the Republic of Uzbekistan [4]. The dataset in question represents population data across multiple consumption sections, spanning 2023 years as shown Figure 1. Each row in the dataset corresponds to a specific consumption section, while the columns represent various types of energy. The numerical values reflect the population counts for each region in a given year.

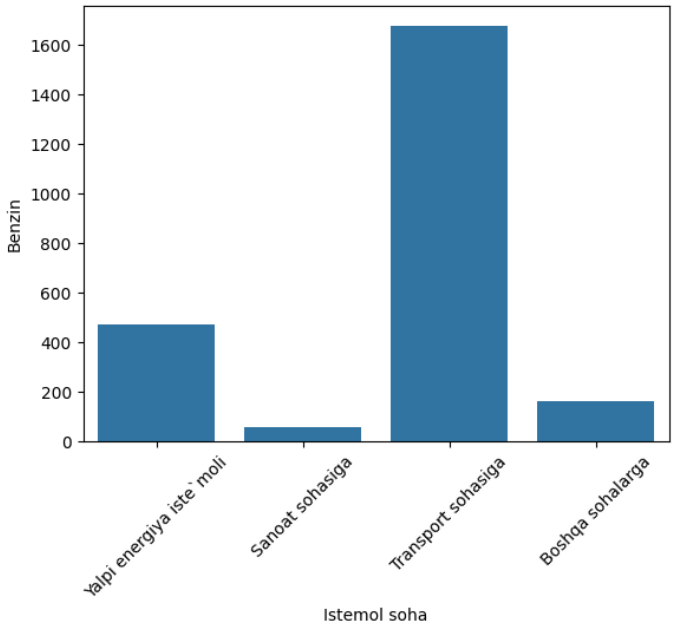
Figure 1. Data description.

```
data_e.head(10)
```

	Istemol soha	Ko`mir	Tabiiy gaz	Neft, gaz kondensatini qo`shgan holda	Benzin	Dizel yoqilg`isi	Mazut	Gazlar suyultirilgan uglevodorod	Kerosin	Koks	Boshqa turdagi yonilg`ilar	Atoma energiyasi	Elektr energiyasi	Issiqlik energiyasi
0	Yalpi energiya iste`moli	3883.4	39091.5	2847.4	473.30	114.6	738.6	1105.7	-84.0	0.1	515.8	-	881.1	-
1	Sanoat sohasiga	427.8	5399.8	-	59.70	169.7	9.9	187.8	0.1	0.3	-	-	2098.7	6.3
2	Transport sohasiga	3.4	4180.3	-	1677.00	801.2	1.2	506.3	177.4	-	-	-	146.3	-
3	Boshqa sohalarga	551.9	11845.0	-	163.72	395.7	140.6	1355.8	41.6	-	-	-	2933.8	2060.9

The bar chart illustrates the consumption of gasoline across various sectors (indicated on the x-axis as "Istemol soha") with the amount of gasoline (labeled as "Benzin") on the y-axis as shown in Figure 2. The Transport sector shows the highest gasoline consumption, significantly more than any other sector, while the Sanoat sohasiga (Industry) has the least consumption. Yalpi energiya iste’moli and Boshqa sohalarga have moderate consumption compared to the other categories.

Figure 2. Petrol consumption bar plot.



This bar chart visualizes the consumption of diesel fuel ("Dizel yoqilg'isi") across different sectors ("Istemol soha"). The x-axis lists the sectors, and the y-axis shows the amount of diesel consumed. The Transport sector consumes the highest amount of diesel fuel, similar to the previous gasoline chart. Boshqa sohalarga (Other sectors) comes second in diesel consumption, showing a significant level of usage compared to other sectors. Both Sanoat sohasiga (Industry) and Yalpi energiya iste’moli (Total energy consumption) have relatively low diesel consumption, with Yalpi energiya iste’moli being the smallest. Overall, the transport sector is the most diesel-dependent, while the energy and industry sectors show minimal reliance on diesel compared to others.

Figure 3. Disel consumption bar plot.

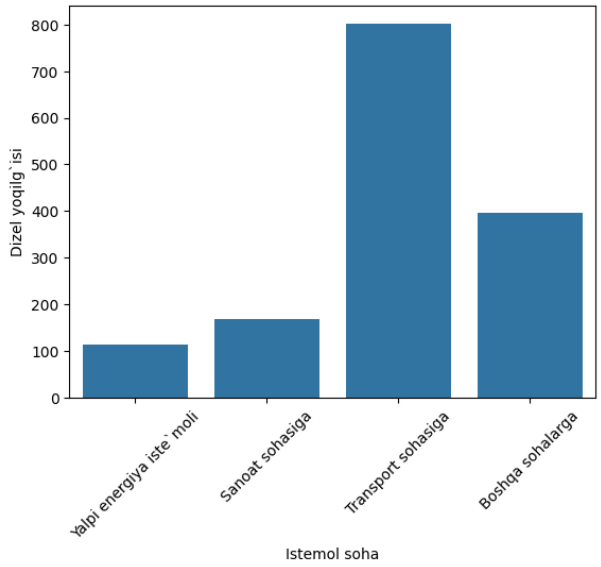
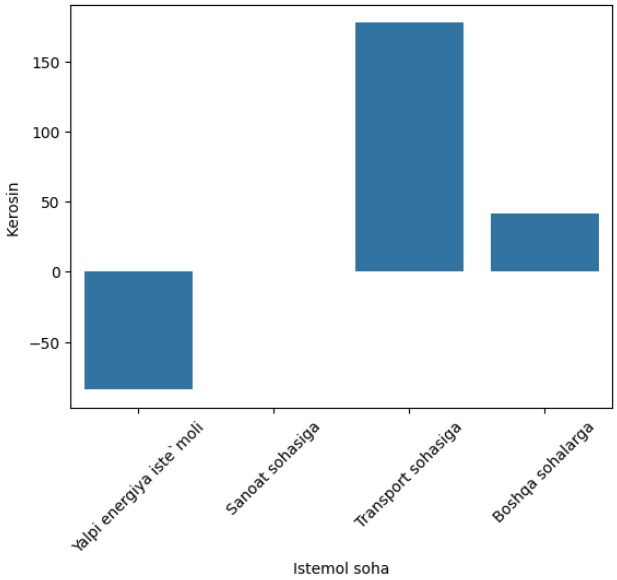


Figure 2 we provided appears to be a bar plot showing population values over the years (likely from

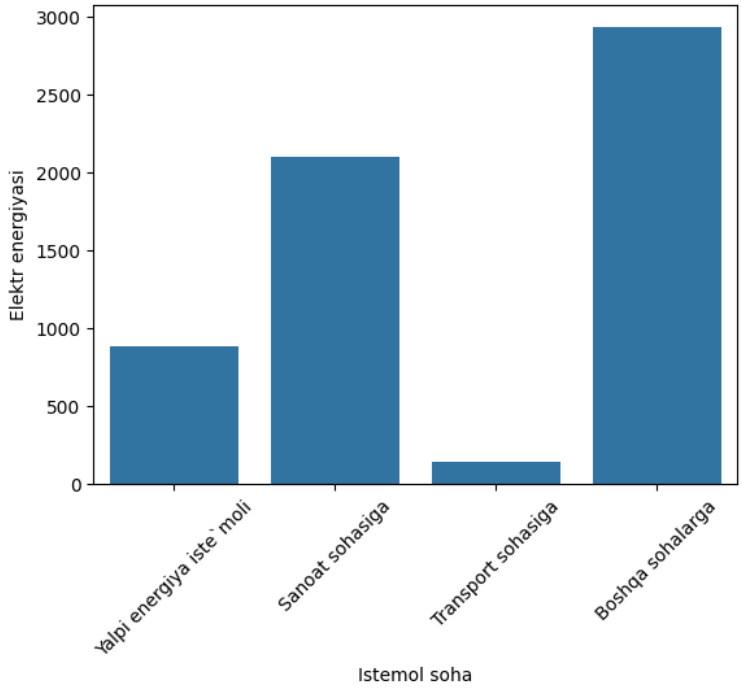
The bar chart displays kerosene consumption ("Kerosin") across various sectors ("Istemol soha") with the x-axis representing the sectors and the y-axis showing the kerosene consumption values as illustrated in Figure 4. The Transport sector once again exhibits the highest consumption of kerosene. The Boshqa sohalarga (Other sectors) shows moderate kerosene consumption. Interestingly, Yalpi energiya iste’moli (Total energy consumption) has a negative value, which might indicate a correction, a return, or an offset for some reason, as negative consumption values typically aren't standard in energy reports. Sanoat sohasiga (Industry sector) shows no visible consumption on the chart, suggesting it either doesn't use kerosene or has negligible usage in comparison to the other sectors. In summary, the Transport sector is a primary consumer of kerosene, while the negative value in the Yalpi energiya iste’moli category raises questions about potential corrections or adjustments in this dataset.

Figure 4. Kerosene consumption.



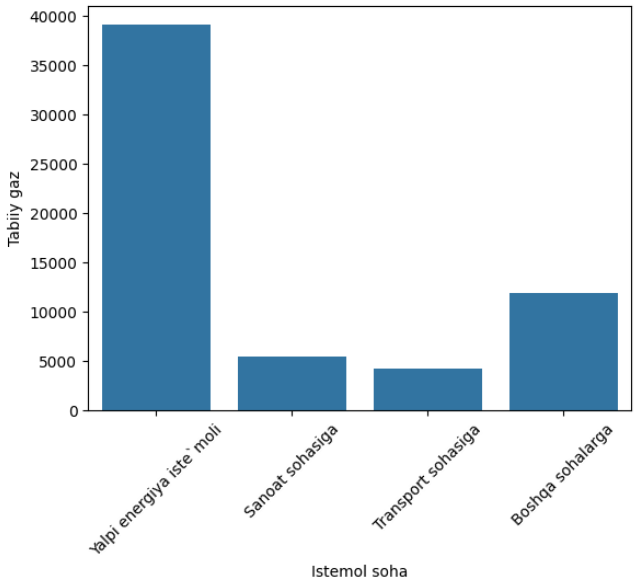
In Figure 5 bar chart illustrates the consumption of electrical energy ("Elektr energiyasi") across various sectors ("Istemol soha"). The x-axis lists the sectors, while the y-axis represents the amount of electrical energy consumed. The Boshqa sohalarga (Other sectors) consumes the highest amount of electrical energy, standing out significantly compared to the other sectors. Sanoat sohasiga (Industry sector) follows, with a substantial amount of electrical energy consumption, though still notably less than Boshqa sohalarga. Yalpi energiya iste’moli (Total energy consumption) shows moderate electrical consumption. The Transport sector has minimal consumption of electrical energy, much lower than any of the other sectors. The Boshqa sohalarga sector is the dominant consumer of electrical energy, while the Transport sector relies minimally on it. This could reflect differences in energy sources or efficiency strategies across these sectors.

Figure 5. Consumption of electrical energy



The bar chart depicts the consumption of natural gas ("Tabiiy gaz") across various sectors ("Istemol soha") as shown Figure 6. The x-axis represents the sectors, while the y-axis shows the amount of natural gas consumed. Yalpi energiya iste’moli (Total energy consumption) has by far the highest consumption of natural gas, dwarfing the other sectors. Boshqa sohalarga (Other sectors) also consumes a significant amount of natural gas, though it is much lower compared to Yalpi energiya iste’moli. Both Sanoat sohasiga (Industry sector) and Transport sohasiga (Transport sector) show relatively low natural gas consumption, with Transport sohasiga having the smallest usage among the sectors. The Yalpi energiya iste’moli sector is the largest consumer of natural gas, indicating that natural gas is a critical energy source in this category. The Boshqa sohalarga also use a substantial amount of natural gas, while the industry and transport sectors have minimal reliance on it.

Figure 6. Consumption of natural gas .



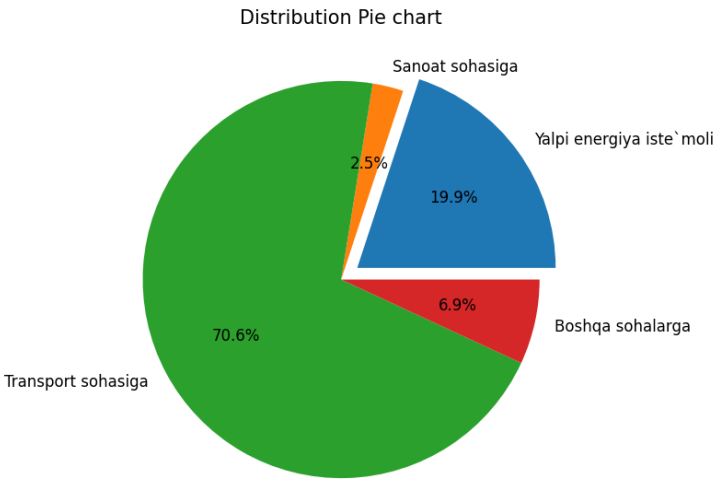
This pie chart represents the distribution of some resource or energy consumption across various sectors as shown Figure 7. Each slice of the pie corresponds to a sector, with the percentage labels showing the share of each sector in the total consumption.

Transport sohasiga (Transport sector) takes up the largest portion, accounting for 70.6% of the total distribution. Yalpi energiya iste`moli (Total energy consumption) is the second-largest sector, consuming 19.9% of the total. Boshqa sohalarga (Other sectors) represents 6.9% of the total. Sanoat sohasiga (Industry sector) has the smallest portion, making up only 2.5% of the total.

Key observations:

- The Transport sector dominates the distribution with more than two-thirds of the total consumption, indicating its significant reliance on this resource.
- The Total energy consumption sector also consumes a notable portion, though it is much smaller compared to the transport sector.
- Other sectors and the Industry sector play minor roles in the overall distribution, with very small shares of the total consumption.

Figure 7. Polulation distribution across regions.



Conclusion

Across the different types of energy and resources (gasoline, diesel, kerosene, electricity, and natural gas), the Transport sector consistently emerges as the dominant consumer, except in the case of electricity and natural gas, where other sectors take the lead. Yalpi energiya iste’moli plays a significant role in overall energy consumption, particularly in natural gas usage. Other sectors such as Sanoat (Industry) and Boshqa (Other) contribute less to the total energy usage but are still relevant in specific cases like electricity and diesel consumption.

Reference:

1. The pandas development team. (2020, February). pandas-dev/pandas: Pandas (latest version) [Software]. Zenodo. https://doi.org/10.5281/zenodo.3509134
2. The Tunisian Company of Electricity and Gas (STEG): https://www.kaggle.com/code/abdulazizmeliboyev/eda-and-modeling-fraud-detection-in-elec/edit
3. Fraud Detection in ELEC and GAZ Consumption: https://www.kaggle.com/code/khsamaha/lightgbm-fraud-detection-in-elec-and-gaz-r
4. Atatistics Agency under the President of the Republic of Uzbekistan: https://stat.uz/uz/
5. World bank: https://www.worldbank.org/en/home
6. United Nations Development Programme (UNDP): https://www.undp.org/
7. Meliboev, A. (2024). IOT NETWORK INTRUSION DETECTION SYSTEM USING MACHINE LEARNING TECHNIQUES. QO‘QON UNIVERSITETI XABARNOMASI, 11(11), 112–115. https://doi.org/10.54613/ku.v11i11.972
8. Meliboev, A. (2024). MACHINE LEARNING METHODS FOR BREAST CANCER CLASSIFICATION BY USING DATA SCIENCE TECHNIQUES. QO‘QON UNIVERSITETI XABARNOMASI, 11(11), 101–104. https://doi.org/10.54613/ku.v11i11.969