# Japan's regional GDP and heatstroke Rate*

Moohaeng Sohn

February 15, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

## Table of contents



# 1 Introduction

Japan is a country located near the edge of eurasia tectonic plates. Numerous earthquakes affect Japan each year. 2011 Tōhoku earthquake was one of the most devastating earthquakes

*Code and data are available at: https://github.com/alexsohn1126/japan_power_and_mortality. Replication aspects in this paper is available at: https://doi.org/10.48152/ssrp-zk6r-9k16

to occur in recorded history. The earthquake was measured at the magnitude X and took over Y number of people's lives [TODO: ADD CITATION HERE]. Not only that, the earthquake was followed by a tsunami, being over Z meters tall and swept over ABC region [TODO: Add actual data here]. The nuclear power plant that was located near the shores of Fukushima prefacture was damaged by the earthquake, followed by the tsunami. The radioactive fuel rods melted and radiation was spread over a wide region near the shores. The International Nuclear Event Scale (INES) intends to measure the seriousness of nuclear accidents. Fukushima nuclear disaster was given INES level 7. Only other nuclear disaster that was given a INES level 7 was the Chernobyl disaster[TODO:ADD CITATION HERE].

Fukushima nuclear disaster instilled fear for nuclear power on the Japanese citizens. The Japanese power companies started shutting down nuclear power plants around the country. This decreased the number of available power throughout the country. A campaign to encourage electricity saving ran over Japan. The campaign was a success, decreasing electricity consumption by X amount [TODO: ADD CITATION]. However Japan has very hot and humid summers which cause heatstrokes especially in elderly populations. Limiting electricity usage by keeping the ACs off could cause higher rate of heatstrokes in vulnerable populations.

Guojun He and Takanao Tanaka published a study which observed the effects of these energy savings and mortality rates due to heatstrokes [TODO: ADD CITATION HERE]. They have found there were higher number of ambulance calls caused by heatstrokes after the shutdown of nuclear power plants. Also, Higher mortality rates in extreme heat (>30°C) was observed after the shutdown. They have dove in depth on how the shift in power generation demographic have possibly affected mortality rates from heatstrokes, but the effect which the cost and the affordability of electricity had was not discussed in detail.

In this paper, I have used the same dataset that was given by Goujun and Takanao to focus on the different age groups of japan, and how they were affected by the change of variables related to electricity affordability. My estimand of interest is to which degree those variables affect the number of heatstroke cases across different age groups in Japan. In this paper, I will first introduce our dataset, going into detail how they were collected by the original authors, and what I have done to clean them if any. Then, I will discuss about the results we have gotten from the data. Lastly, I will discuss what we have found, and their possible implications. I will be reflecting upon our process and discussing potential future researches.

# 2 Data

## 2.1 Sources

The original paper by Goujun and Takanao combines datasets from multiple sources. We will go through where each variable was sourced from.

The electricity saving target dataset from 2008 to 2015 was collected from Japan's Electricity Supply-Demand Verification Subcommittee for every summer season [CITE ORIGINAL PAPER]. The paper defined the summer season as July, August, and September. We will also use the same definition of summer in our paper.

Our dataset is organized by "areas". Areas are a collection of prefectures (states) in Japan which have the same power companies. There are 10 major power companies within Japan, therefore we divide Japan into 10 areas. This is fine because each power company essentially holds a monopoly over their region [CITE ORIGINAL PAPER]. heatstroke data from 2008 to 2015 was collected by observing the number of ambulance transports caused by a heatstroke. The heatstroke data was then converted into the heatstroke rate per 100,000 people. This data was gathered from the Fire and Disaster Management Agency in Japan [CITE ORIGINAL PAPER]. This data is collected in a monthly basis. The prefectural temperature data was collected by consulting Meteorological Agency of Japan [CITE ORIGINAL PAPER]. This was done by aggregating station level data. Original authors weighed each station data with the inverse square of the distance to the population center of a prefecture. This meant that weather stations closer to the prefecture's population center would be factored in more than the stations further away.

GDP per capita for each area was computed by taking the sum of GDP from all the prefectures in each area, and dividing it by the number of people in the area. Number of people in the area was computed by adding up the number of people in the prefectures within the area.

We downloaded Japan's map divided into prefectures to help us visualize data better. This was done by using the R package `rnaturalearth`. This R package downloads data from a website called Natural Earth (Kelso and Patterson, n.d.). This website hosts public domain map data, made by volunteers around the world.

We chose this dataset because it is well documented by the original paper authors. It also combines data from numerous sources, making it easier for us to clean up and combine the data into one table. As of our knowledge, there isn't a dataset which has electricity and heatstroke rate of Japan in the same dataset.

## 2.2 Methodology

To combine many datasets into one, we have used Open Source Statistical Programming Language R (R Core Team 2022). Along with it, we have used various R packages to assist us in downloading and processing the data, and generating graphs. These packages are: `tidyverse` (Wickham et al. 2019), `rnaturalearth` (Massicotte and South 2023), `sf` (Pebesma 2018), `here` (Müller 2020), `patchwork` (Pedersen 2024), and `haven` (Wickham, Miller, and Smith 2023).

The original paper's authors have used the Stata statistical software. Meaning the data files were given in a `.dta` file format. Using `haven` R package, we were able to load it into R and work with the dataset.

## 2.3 Variables Summary

Figure 1 shows our 5 main variables of concern: electricity price, heatstroke rate, electricity saving target, average temperature, and GDP per capita (for each area). Electricity price tends to have higher prices after 2011. It's lowest price was achieved in 2010, and highest was achieved in 2014. heatstroke rate doesn't show a very strong trend but it does seem like 2010 had the highest heatstroke rate, and lowest in 2009. Electricity saving target did not really take off before the Fukushima disaster. We observe that the maximum was achieve in 2011, with minimums being 0 across all years before 2011. There doesn't seem to be a clear trend in average temperature over the years. GDP per capita for each power company does seem to have somewhat of a dip then a recover by the year 2015.

Table 1 summarizes these key variables. One notable thing is that the standard deviation for GDP per Capita is 200, which indicates most of the data is centered relatively close around the mean, or the average. Standard deviation measures how spread out the data is. Higher standard deviation means data points are spread far apart. Another notable thing is that the maximum heatstroke rate is 71 while mean is 12 and standard deviation is 12.27.

There are total 3384 data points in the dataset. Each row of the dataset can be identified by the year, month, area, and the age group. This means that variable such as GDP stays constant for the same year in the same area.

## 2.4 Measurements

The mean temperature data is measured in Celsius, and is measured by 1,300 weather stations scattered throughout Japan, and the results are aggregated and weighed based on the distance away from the population center of the prefecture (He and Tanaka 2023). The electricity price is measured in yen per Kwh. The paper does not specify how this was measured. One possibility is they take the revenue generated by each electric company and divide it by the power generated by the company. The heatstroke rate is a number which represents how many ambulances were used due to a heatstroke per 100,000 people. This is divided into age groups as well. It is not mentioned what kind of criteria must be met for an ambulance use to be considered used due to a heatstroke. Electricity saving target rate is a decimal point number which represents what the power companies and the Japanese government aimed to decrease electricity consumption by. This is an arbitrary target set by the Japanese Electricity Supply-Demand Verification Subcommittee in the summer every year (He and Tanaka 2023). They factor in many variables. It is not documented this arbitrary number was chosen.

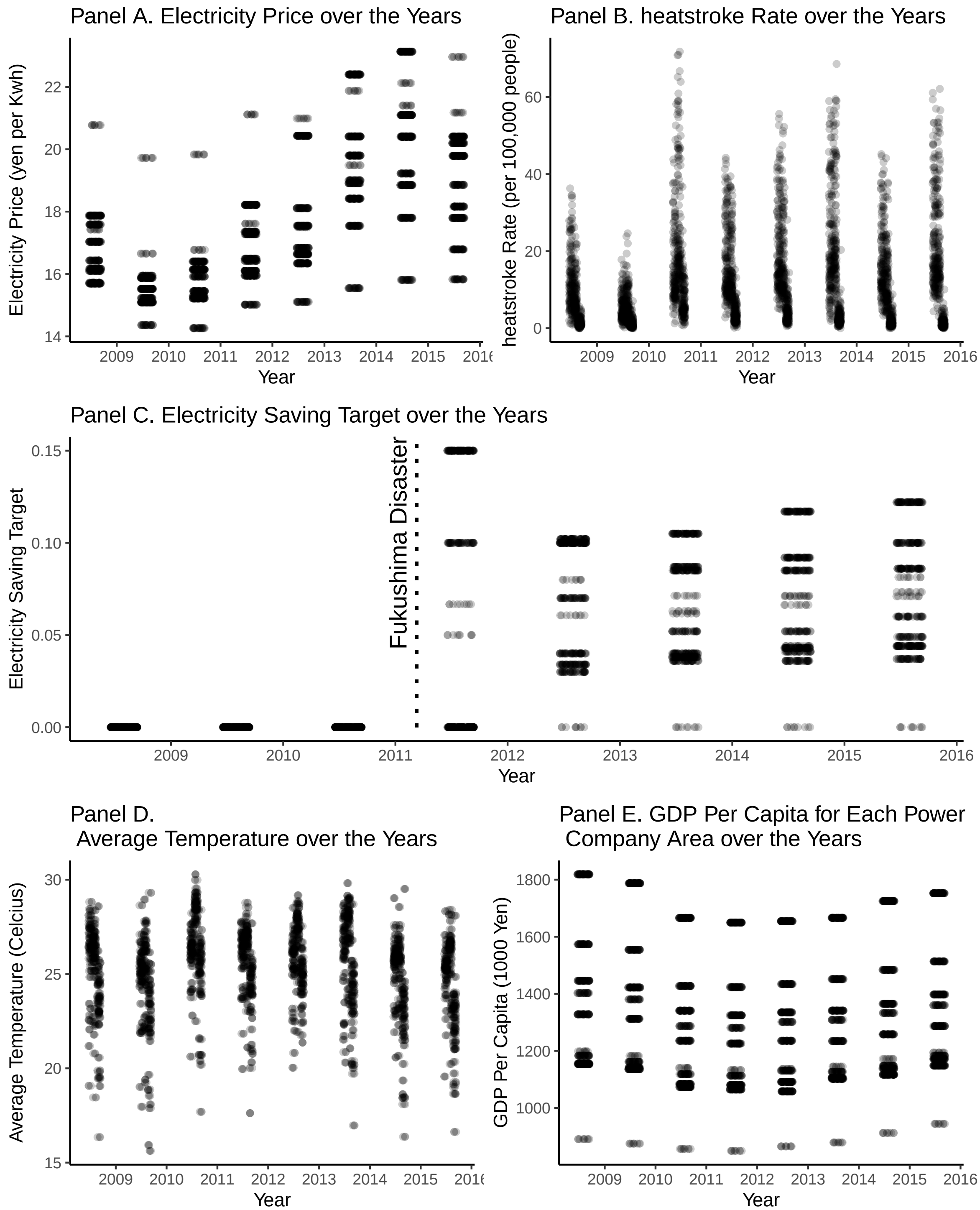


Figure 1: Summary of Key Variables for Each Year

Table 1: Summary statistics of key variables

| Variables | Max | Mean | Median | Min | Standard Deviation |
|---|---|---|---|---|---|
| Electricity Saving Target | 0.150 | 0.042 | 0.036 | 0.000 | 0.046 |
| heatstroke Rate | 71.794 | 12.476 | 9.068 | 0.000 | 12.270 |
| Electricity Price | 23.130 | 17.637 | 17.153 | 14.264 | 2.126 |
| Average Temperature | 30.288 | 25.152 | 25.620 | 15.624 | 2.370 |
| GDP per Capita | 1818.833 | 1317.194 | 1283.925 | 849.792 | 227.539 |

## 3 Results

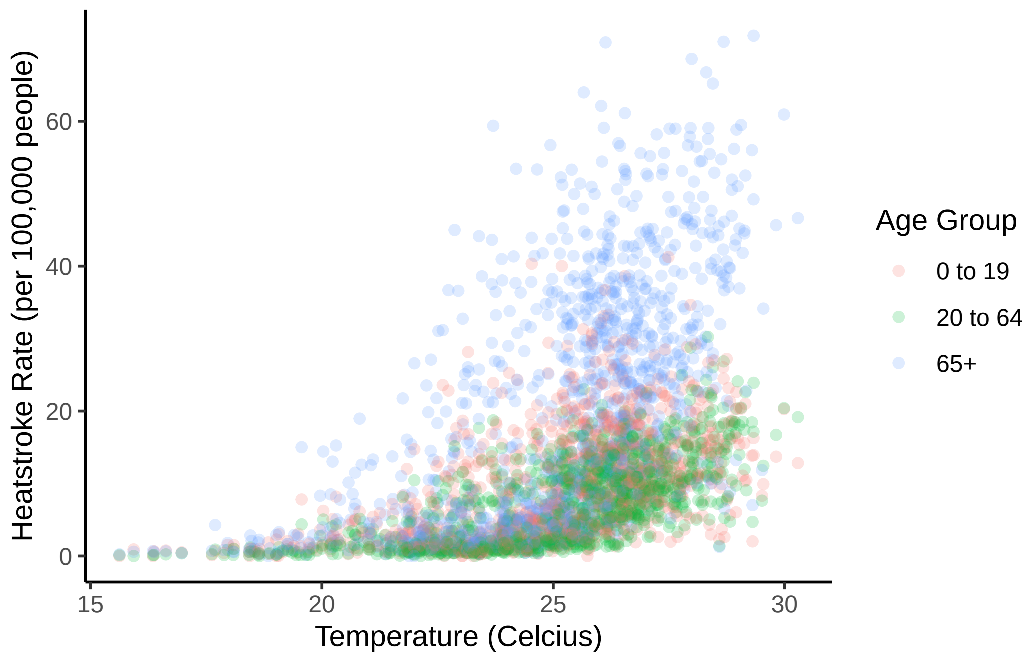


Figure 2: Temperature and Heatstroke Rate by Age Group

We now explore how these variables interact with each other. We start with perhaps the most obvious interaction: the mean temperature and the heatstroke rate. Figure 2 shows a scatter plot demonstrating that as temperature increases, the heatstroke rate increases. An interesting point is how these trends seem to differ significantly between different age groups. Those who are 65 or older have noticeably higher rate of heatstroke at the same temperature.

We can divide the GDP per capita by the electricity prices to compute how affordable the electricity is for a person working and living in that area. It will be a measurement of how much Kwh of energy can be bought with GDP per capita for different areas. Figure 3 shows that the prefectures in the center of japan has the most affordable electricity. Not only that, we can also look at how the electricity affordability have changed throughout the years by looking at Figure 4. It can be seen that prior to 2010, electricity was more affordable overall, but after 2011,

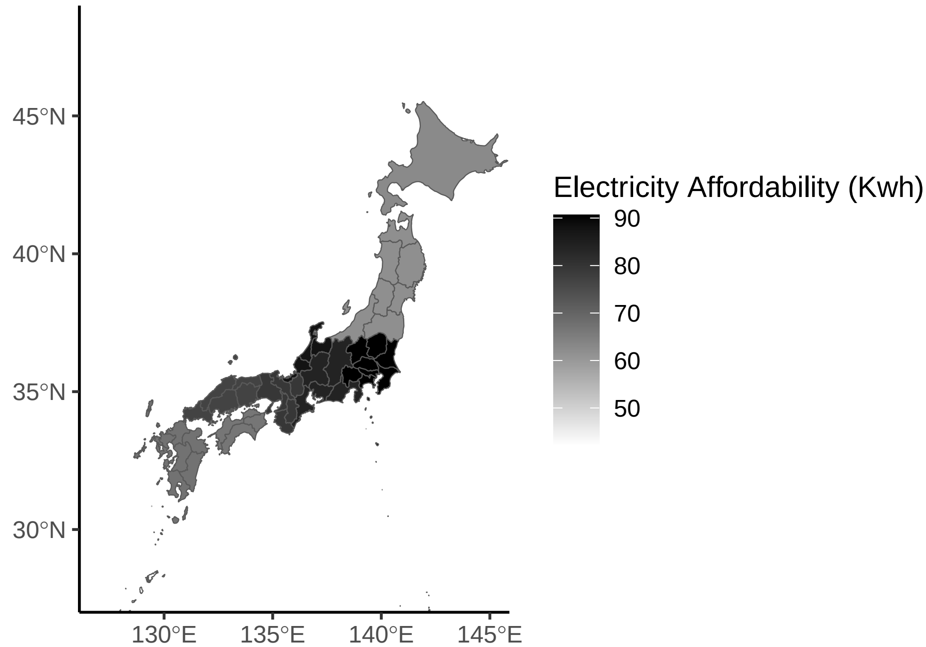


Figure 3: Electricity affordability around regions of Japan

# 4 Discussion

## 4.1 First discussion point

## 4.2 Second discussion point

## 4.3 Third discussion point

## 4.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

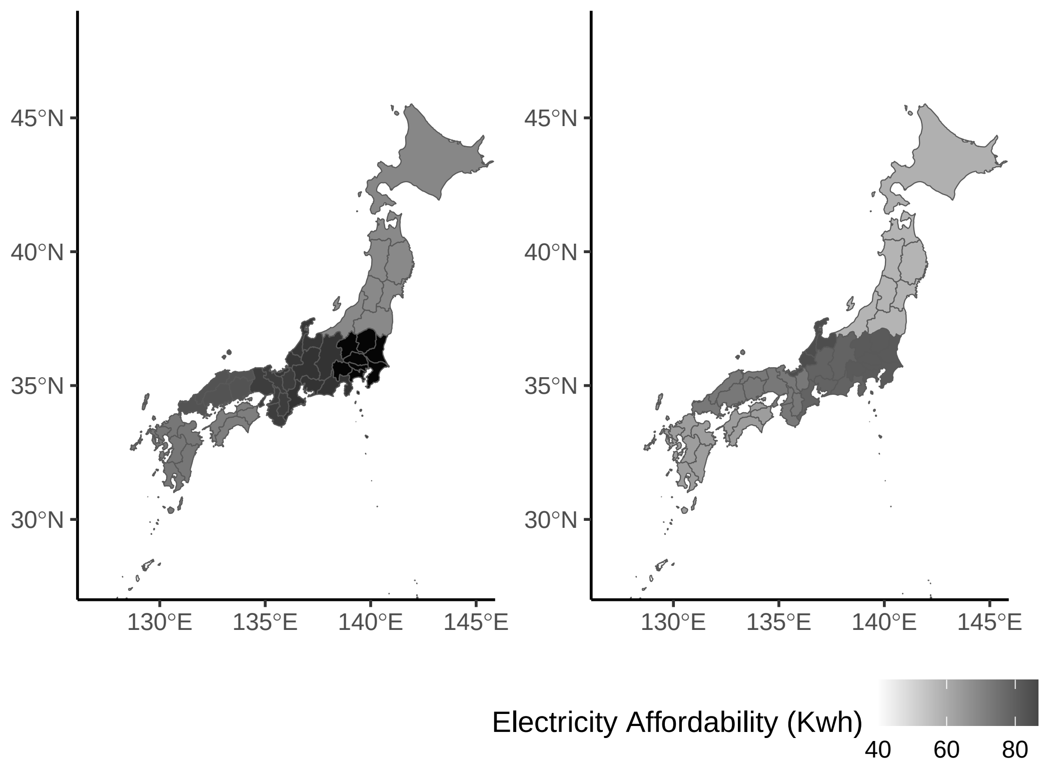


Figure 4: Electricity affordability around regions of Japan from 2008 to 2010 (Left), and 2011 to 2015 (Right)

## References

He, Guojun, and Takanao Tanaka. 2023. "Energy Saving May Kill: Evidence from the Fukushima Nuclear Accident." *American Economic Journal: Applied Economics* 15 (2): 377–414. https://doi.org/10.1257/app.20200505.

Kelso, Nathaniel Vaughn, and Tom Patterson. n.d. *Natural Earth*. https://www.naturalearthdata.com/about/.

Massicotte, Philippe, and Andy South. 2023. *Rnaturalearth: World Map Data from Natural Earth*. https://CRAN.R-project.org/package=rnaturalearth.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.

Pedersen, Thomas Lin. 2024. *Patchwork: The Composer of Plots*. https://CRAN.R-project.org/package=patchwork.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Evan Miller, and Danny Smith. 2023. *Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. https://CRAN.R-project.org/package=haven.