International Journal of Cyber-Security and Digital Forensics (IJCSDF) 7(3): 200-208
The Society of Digital Information and Wireless Communications (SDIWC), 2018 ISSN: 2305-001

Survey and Analysis of Regional Characteristics of Unmanaged Stray IoT devices

Yuki Nakazawa
Tokyo Denki University
Senju-Asahi-cho, Adachi-ku, Tokyo
Email: nakazawa@isl.im.dendai.ac.jp

Ryoichi Sasaki
Tokyo Denki University
Senju-Asahi-cho, Adachi-ku, Tokyo
Email: r.sasaki@mail.dendai.ac.jp

Atsuo Inomata
Tokyo Denki University
Senju-Asahi-cho, Adachi-ku, Tokyo
Email: inomata@mail.dendai.ac.jp

ABSTRACT

Recently, various devices, such as automobiles, and medical and industrial systems, have been connected to the Internet, and this network is referred to as the Internet of Things (IoT). IoT devices that do not have an authentication function risk unauthorized access. In addition, there is a high possibility that the IoTs may be used as a springboard for Distributed Denial of Service (DDoS) attacks. Moreover, once the installed equipment is no longer used, it is difficult to notice attacks, such as illegal operation by a malicious user, so that the risk increases further. These unmanaged devices are referred to as stray IoT devices. In this paper, we describe the status of stray IoT devices in Japan and report that we investigated the relationship of regional characteristics between the number of stray IoT devices by prefecture and other data by prefecture by conducting a correlation analysis. As a result, a significant correlation between the phishing damage rate and the number of stray IoT devices was found to exist. These results are thought to be useful for planning countermeasures against stray IoT devices.

KEYWORDS

Internet of Things, IoT Security, Stray IoT, Regional characteristics, Correlation analysis

1 INTRODUCTION

Recently, various devices, such as automobiles and medical and industrial systems, have been connected to the Internet, and this network of devices is referred to as the Internet of Things (IoT). The number of IoT devices is increasing dramatically, and, according to the White Paper on Information and Communication [1], is estimated to reach 53 billion by 2020.

However, many IoT devices that do not have an authentication function can be operated freely by malicious individuals over the Internet using a browser. The same problem may occur if the authentication function is not set correctly. In addition, there is a high possibility that IoT devices may be used as a springboard for DDoS attacks.

Moreover, if the installed device does not have an authentication function and is no longer being managed, attacks become difficult to notice, and the risk increases further. We refer to these unmanaged stray IoT devices.

In this paper, we describe a survey on the actual situation of Stray IoT devices in Japan for considering countermeasures. As a countermeasure against the problem of IoT devices becoming stray IoT devices, we considered users of IoT devices. We found that more effective countermeasures were based on the temperament and characteristics of residents of various prefectures. Therefore, we carried out a correlation analysis between the temperament of the residents of prefectures and the regional characteristics of prefectures and the data on stray IoT devices investigated by prefecture in Japan. To our knowledge, this is the first analysis of the regional characteristics of stray IoT devices.

2 RELATED RESEARCH

A number of studies have investigated devices infected by malware on the IoT. Suzuki et al. observed attacks using honeypots disguised as IoT devices. [2] Then, by conducting device fingerprinting on the observed host, the device at the source of the attack was estimated. As a

result, 15,096 devices corresponding to 14% of observed hosts were found to be IoT devices such as DVR, router, and IP camera.

Eric explored how botnets in smart devices are exacerbating identity crime and also placed a refrigerator at the heart of IoT that has become connected through the Internet and thereby susceptible to botnets and the collection of personal identification information as an enabler for identity crime. [3]

Kasama et al. classified the source equipment used for attacks by sending port scan packets to TCP port 23 of the IP address obtained by observing darknets having 200,000 IP addresses. [4] As a result, it was possible to collect the banner information of the attack source with 40,973 IP addresses, corresponding to 20.5% of the observation hosts, and IoT devices, such as DVRs and routers, were observed.

Although all of these studies have been useful, they investigated IoT devices that may already be infected by some type of malware. However, there is has been no actual survey of stray IoT devices on the Internet, and there has been no mention of the regional characteristics of stray IoT devices.

3 Proposed method for classification

3.1 Outline of classification

Existing related studies have clarified the risk of IoT devices becoming a springboard for attack by honeypot and darknet observation, but an actual situation survey was not performed.

Therefore, it is important to investigate the actual situation of the stray IoT. Furthermore, we investigated whether the IP network cameras have an authentication setting. The reason for selecting network cameras for the investigation is that network cameras are considered to have a high probability of becoming stray IoT devices, because the main use of network cameras is to provide evidence in the case of an emergency, and, ordinarily, they are left unmanaged.

As a countermeasure against these problems, users of IoT devices are usually warned of the potential for exploitation.

We believe that the characteristics of prefectures in Japan can be used to enhance the effect of such warnings. This is because weather and disaster warnings are presented by prefecture and region, and users' interests in local information tend to be higher.

Therefore, we estimated the number of target devices by prefecture and analyzed the correlation with other data by prefecture.

3.2 Method of investigation

The method of investigation is described in three steps as follows.

[Step 1] Investigation of the existence of an authentication setting

We investigated the status of an authentication setting by sending an HTTP request for the IP addresses of IoT devices.

[Step 2] Classification of prefectures in which web cameras are installed

[Step 2-1]

The Classification was performed using the characteristic that the prefecture name or code is part of the domain.

Moreover, the domain was acquired by reverse lookup by IP address with the nslookup command.

[Step 2-2]

Since the data obtained in Step 2-1 may be biased due to the population of the prefecture, we use a modified number, i.e., the number of devices per 100,000 residents.

[Step 3] Correlation analysis

Correlation analysis between the results obtained in Step 2-2 and various prefecture-specific data is performed.

The details of the data set used for the survey are shown in the next section.

3.3 Data Set

We used Censys [5] to collect the data set with respect to the IP address of the IoT devices. Censys is a novel search engine provided by the research team at the University of Michigan.

We selected 9,999 IP addresses of IoT devices that satisfy the following three criteria:

- The country code is Japan.
- A web camera is included in the page title.
- An OCN provider is used for the Web camera communication.

The number of IoT devices is set to be 9,999 because this is the maximum value permitted by the Censys provider. In the present study, we consider IoT devices using an OCN provider because percentage of OCN providers among total users is 34.5%, which is the highest in Japan.

4 EXPERIMENT

4.1 Investigation of Connection Status

We executed and collected the connection state by HTTP request. The results of the experiment are shown in Table 1.

Table 1. Results of investigation of connection status

Connection status	Number of hosts
Connectable without authentication	477
Connectable with authentication	1,530
No reply	7,992

As a result, we found that a total of 477 (5%) hosts can connect without authentication. Moreover, 1,530 (15%) hosts were found to be connectable following authentication.

Moreover, we connected to the host in the connectable state using a browser. An example is shown in Figure 1, in which a mosaic process was applied to the actual page image for privacy protection.

Figure 1 Example of a connectable host.

If the IP address of the connectable state host can be determined, anyone can connect to the operation page of the network camera, as shown in Figure 1.

Figure 1 shows an image of a video that was taken by an IP network camera. An operation panel is also shown on the left-hand side of the screen, through which it is possible to pan and tilt, as well as zoom, the network camera. Furthermore, it is possible to change the resolution, image quality, and display size of the captured video image.

Network cameras are installed in various places, such as inside factories, offices, shops, and foyers, as well as on eaves and in garages. Network cameras installed in factories, offices, and shops were originally installed for the purposes of crime prevention, such as monitoring restricted areas or preventing shoplifting. However, the installation of the network camera leads to a reduction in the crime prevention level, which may increase crime risk. If video images are released, there is a risk that important documents related to employees or the trade secrets will be revealed.

Moreover, by manipulating the orientation of network cameras placed in foyers or under eaves, personal information, such as the names and addresses of the users, may be revealed. Behaviors such as residents leaving their houses may be monitored, which may lead to crimes such as burglary.

The current state, in which there are many network cameras without authentication, is very dangerous. Moreover, the results of the present investigation and experimentation have revealed that the authentication settings of the administrator page have not been used in four of the 477 connectable network cameras. An example of an administrator page is shown in Figure 2.

Figure 2 Example of an administrator page.

The administrator page shown in Figure 2 has buttons that can be used to, for example, change the network settings, create an administrator, and change the password. Moreover, the user can confirm the connection state, change the color of the indicator lamp on the main body of the IP network camera, reboot the camera, and change the factory state value. Furthermore, information such as the MAC address of the corresponding terminal, the firmware version, and protocols, can be viewed. In particular, in changing the authentication settings, a malicious user may hijack the camera.

Moreover, in recent years, in order to reduce resource usage and cost, paper manuals are no longer provided and have been replaced with electronic manuals, which are published on the manufacturer's or vendor's website.

As such, the user ID and default password can be determined by anyone, and for devices that have default authentication settings, the password must be changed. In the present study, we also checked for the presence or absence of authentication settings. However, we did not investigate whether the default password was used in the terminal used for authentication because of concerns that doing so may violate the illegal access prohibition law. A certain number of terminals are presumed to use the default password, which is extremely dangerous depending on the setting, even for the terminal for which authentication is performed.

4.2 Classification of Prefectures in Japan

Through reverse lookup of internet domain against the data set, we identified the prefectures of 9,225 hosts. Five prefectures with a large number of hosts and five prefectures with few hosts, together with the population ranking [6], are shown in Tables 2 and 3, respectively.

Table 2 Prefectures with a large number of hosts

Prefecture	Total number of hosts	Population ranking
Tokyo	1,837	1
Osaka	1,507	3
Aichi	704	4
Hokkaido	443	8
Kanagawa	342	2

Table 3 Prefectures with few hosts

Prefecture	Total number of hosts	Population ranking
Shimane	34	46
Wakayama	33	40
Kagawa	30	39
Tokushima	30	44
Saga	15	42

Tokyo had the largest number of hosts (1,837), followed by Osaka, Aichi, Hokkaido, and Kanagawa. Saga Prefecture had the fewest hosts (15). Based on these results, the number of hosts is proportional to the population. In other words, prefectures with high populations have many hosts.

In addition, five prefectures with a large number of connectable hosts are listed Table 4, together with the corresponding population rankings.

Table 4 Prefectures with a large number of connectable hosts

Prefecture	Number of connectable hosts	Population ranking
Tokyo	45	1
Hokkaido	25	8
Osaka	20	3
Toyama	18	37
Aichi	17	4

Of the five prefectures in which there were a large number of hosts in the connectable state shown in Table 4, four prefectures (Tokyo, Hokkaido, Osaka, and Aichi) were also the prefectures with a high number of hosts, as shown in Table 2. However, in Toyama prefecture, despite having a population ranking of 37th (out of 47), the number of connectable hosts is 18, which is high.

4.3 Number of connectable units per 100,000 people

As shown in Section 4.2, the number of hosts in a connectable state is proportional to the population. Therefore, we investigated the number of connectable states per 100,000 residents. Five prefectures with a large number of connectable hosts per 100,000 residents are listed in Table 5, together with the corresponding population ranking. In addition, five prefectures with a small number of connectable hosts per 100,000 residents are listed in Table 6, together with the corresponding population ranking.

The results of the survey are shown in Appendix 1.

Table 5 Prefectures with many connectable hosts per 100,000 residents

Prefecture	Number of connectable hosts per 100,000 residents	Population ranking
Toyama	1.682243	37
Kagawa	0.815494	39
Niigata	0.648508	15
Iwate	0.623052	32
Tottori	0.522648	46

Table 6 Prefectures with few connectable hosts per 100,000 residents

Prefecture	Number of connectable hosts per 100,000 residents	Population ranking
Miyazaki	0.089767	36
Ishikawa	0.086505	34
Saitama	0.082884	5
Nara	0.072674	30
Okinawa	0.070373	25

As shown in Table 5, Toyama Prefecture had the largest number of connectable hosts per 100,000 residents, more than twice that of Kagawa Prefecture, which had the next largest number of connectable hosts.

4.4 Survey of the Ratio of Connectable Hosts

We investigated the ratio of connectable hosts for each prefecture based on a survey of connection state and the results of prefecture classification. Five prefectures having a high percentage of connectable hosts are listed in Table 7, together with the corresponding population ranking.

Table 7 Prefectures with a high percentage of connectable hosts

Prefecture	Percentage of connectable hosts	Population ranking
Toyama	66.6	37
Kagawa	66.6	39
Iwate	66.6	32
Tottori	60	46
Shimane	60	46

As shown in Table 7, many prefectures having a high percentage were also listed in Table 5. Based on this result, the risks of Toyama and Kagawa prefecture are clearer. Shimane Prefecture was ranked 12th with respect to the number of connectable hosts per 100,000 residents.

4.5 Correlation analysis

We performed a correlation analysis on the number of connectable units per 100,000 residents, which was investigated in Section 4.3, and various data by prefecture. Table 8 shows the correlation coefficient, p-value, and significance of data for which significant correlation was obtained.

Table 8 Correlation analysis results of connectable hosts per 100,000 residents

Data by prefecture	Correlation coefficient	p-value	Significance
Annual rainfall [7]	0.33678	0.021	Strongly significant
Annual snowfall days [7]	0.31374	0.032	Strongly significant
Number of elderly per 100,000 residents [8]	0.27404	0.062	Weakly Significant
Phishing damage rate [9]	0.28574	0.052	Significant

The correlation graph with the annual precipitation days is shown in Figure 3, and the correlation graph with the annual snowfall days is shown in Figure 4.

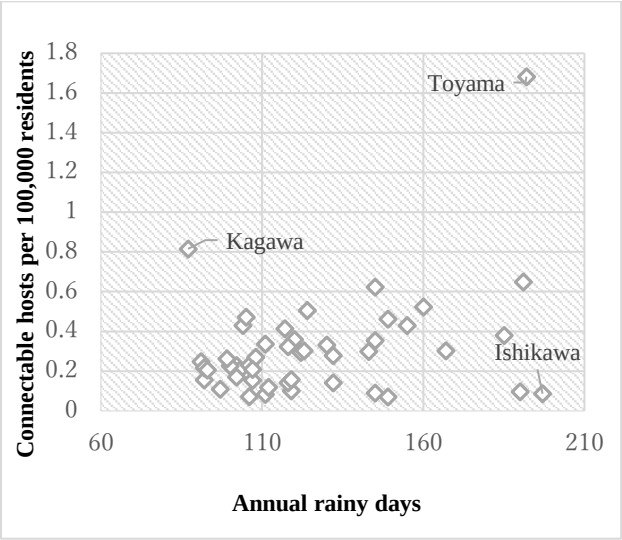

Figure 3 Correlation with annual precipitation days.

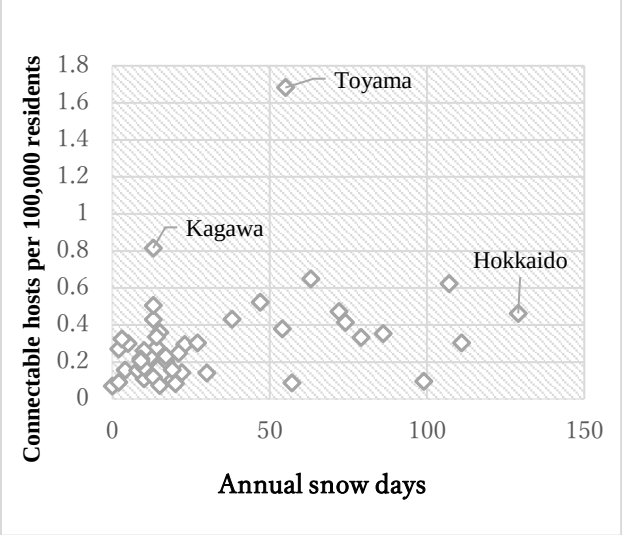

Figure 4 Correlation with annual snow days.

In the correlation with the annual precipitation days shown in Figure 3, the correlation coefficient is 0.33678, which is a positive correlation, and the p-value of 0.021 indicates that the result is significant. In the correlation with the annual snowfall days shown in Figure 4, the correlation coefficient is 0.31374, which is a positive correlation, and the p-value of 0.032 indicates that the result is significant. Thus, the proportion of hosts in the connectable state without authentication is high in the Hokuriku region, which has many precipitation days, and

in the Kyushu region, which has many snowy days.

The correlation graph for the number of elderly per 100,000 residents is shown in Figure 5.

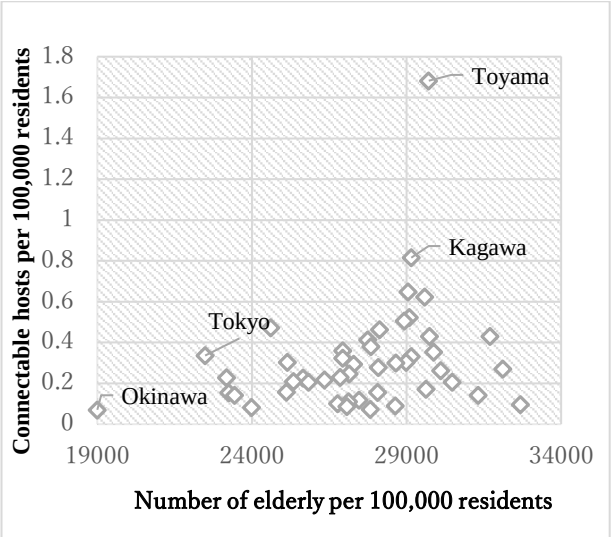

Figure 5 Correlation with the number of elderly per 100,000 residents.

In the correlation with the number of elderly per 100,000 residents shown in Figure 5, the correlation coefficient is 0.27404, which is a positive correlation, and the p-value of 0.062 indicates that the result is weakly significant. Prefectures with large elderly populations are often located some distance from large cities. As such, we believe that there is a tendency for high risk in prefectures having a high ratio of elderly residents because such regions have a shortage of information security experts and people with a high degree of information literacy. The correlation graph with the phishing damage rate is shown in Figure 6.

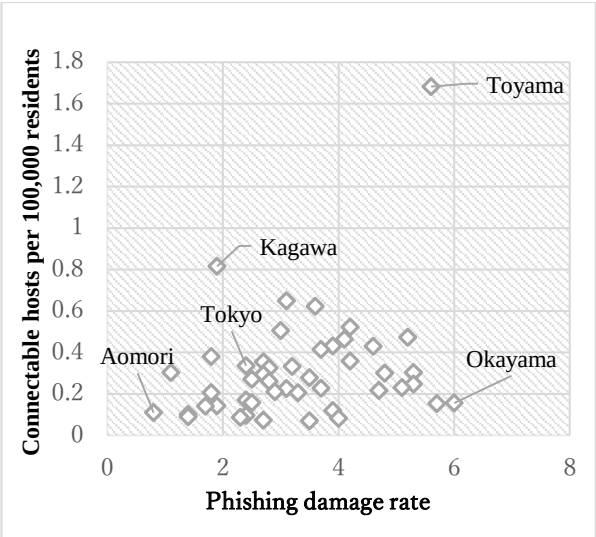

Figure 6 Correlation with phishing damage rate.

In the correlation with the phishing damage rate per 100,000 residents shown in Figure 6, the correlation coefficient is 0.28574, which is a positive correlation, and the p-value of 0.052 indicates that the result is significant. Phishing is a kind of social engineering attack in which malicious users send an email impersonating another user to steal important personal information, such as credit card numbers and account information, such as user IDs and passwords. Social engineering attacks take advantage of victim psychological gaps and behavior. Therefore, since the correlation with the phishing damage rate is a positive correlation, it is conceivable that the daily interest in information security has an influence on the presence or absence of authentication settings in network cameras.

Furthermore, in Section 4.4, we performed a correlation analysis on the percentage of connectable hosts, the number of elderly per 100,000 residents, and the phishing damage rate. The results are shown in Table 9.

Table 9 Correlation analysis results for the ratio of connectable state

Data by prefecture	Correlation coefficient	p-value	Significance
Number of elderly per 100,000 residents [8]	0.38529	0.007	Significant
Phishing damage rate [9]	0.28913	0.048	Significant

As shown in Table 9, the correlation coefficient increased, and the p-value became small, proving the validity of the correlation analysis.

5 CONCLUSIONS

In this paper, we described survey on the connection status of IP network cameras, which are highly likely to be stray IoT devices. As a result, 477 hosts were able to be accessed and operated without authentication, and four hosts

were able to be accessed with administrator privileges. The existence of such devices requires the user to correctly recognize the security risk as the number of IoT devices increases. We believe that it is important to pay attention to users. According to the prefecture classification and the investigation of the number of connectable hosts per 100,000 residents, Toyama prefecture, despite having a population ranking of 37th (out of 47), the number of connectable hosts is 18 revealed to be particularly noteworthy. Furthermore, the risks in Tohoku and Hokuriku regions were found to be high. Based on this correlation analysis, the correlation between the number of elderly per 100,000 residents and the phishing damage rate clearly shows that regional characteristics, such as the awareness of daily information security and concerns, affect the presence or absence of authentication settings.

Thus, we believe that the characteristics of prefectures in Japan can be used to enhance the effect of security warnings.

In the future, we intend to perform a larger-scale survey of stray IoTs other than IP network cameras, and we would like to develop better countermeasures based on a correlation analysis with additional prefectural data that we performed in December of 2016.

The number of hosts in the connectable state, which was 477 units as of December, 2016, declined to 134 units as of April, 2017. Moreover, we found that the number of hosts with authentication increased from 1,530 to 1,717. Although these changes are a good thing, the current situation, in which network cameras, which, by nature, should be private are in fact connectable, is extremely dangerous, and we would like to continue our investigation in the near future.

REFERENCES

1. Ministry of Internal Affairs and Communications Heisei era 27-year version information communication white paper, Information, http://www.soumu.go.jp/johotsusintokei/whitepaper/h27.html
2. S.Suzuki, Yin Minn Pa Pa, Y.Ezawa, Ying,Ying, K.Nakayama, K.Yoshioka T.Matsumoto, ” Enhancement of honeypot IoTPOT observing attacks on embedded devices”, The Institute of Electronics Information and Communication, 2015, ICSS2015-47(2016)
3. Eric Holm, “The Role of the Refigerator in Identity Crime?”, International Journal of Cyber-Security and Digital Forensics. [Online]. 5(1), pp. 1-9. (2016)
4. T. Kasama, S. Shimamura, D. Inoue, ” Detection of attack activity status of embedded device combing passive observation and active observation”, The Institute of Electronics Information and Communication, A Vol.J99-A No.2 pp.94-105, (2016)
5. Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey,J.Alex Halderman, ” A Search Engine Backed by{I}nternet-Wide Scanning”, Proceedings of the 22nd ACM Conference on Computer and Communications Security, (2015)
6. Ministry of Internal Affairs and Communications Statistics Bureau, Population by prefecture and population change rate, Information, http://www.stat.go.jp/data/nihon/02.html
7. Prefectural data ranking http://uub.jp/pdr/
8. Cabinet Office, Heisei era 28-year old society white paper Information, http://www8.cao.go.jp/kourei/whitepaper/index-w.html
9. Ministry of Internal Affairs and Communications Statistics Bureau, Heisei ear 28-year old Servey on trend of communication use, Information, https://www.e-stat.go.jp/

Appendix 1. Survey results

Prefecture	Population (1,000)	Total number of Connectable IP network camera	Authentication setting available	No authentication setting	Percentage of authentication setting	Ratio of elderly residents	Phishing damage rate (%)
Toyama	1,070	27	9	18	1.682242991	29719.62617	5.6
Kagawa	981	12	4	8	0.815494393	29153.92457	1.9
Niigata	2,313	38	23	15	0.648508431	29053.17769	3.1
Iwate	1,284	12	4	8	0.62305296	29595.01558	3.6
Tottori	574	5	2	3	0.522648084	29094.07666	4.2
Nagasaki	1,386	19	12	7	0.505050505	28932.17893	3
Miyagi	2,328	38	27	11	0.472508591	24613.40206	5.2
Hokkaido	5,400	66	41	25	0.462962963	28129.62963	4.1
Shimane	697	5	2	3	0.430416069	31707.31707	3.9
Ehime	1,395	17	11	6	0.430107527	29749.10394	4.6
Fukushima	1,935	18	10	8	0.413436693	27751.93798	3.7
Fukui	790	16	13	3	0.379746835	27848.10127	1.8
Saga	835	6	3	3	0.359281437	26946.10778	4.2
Yamagata	1,131	21	17	4	0.353669319	29885.05747	2.7
Tokyo	13,390	231	186	45	0.336071695	22486.93055	2.4
Nagano	2,109	22	15	7	0.331910858	29160.73969	3.2
Shizuoka	3,705	94	82	12	0.32388664	26936.5722	2.8
Tochigi	1,980	37	31	6	0.303030303	25151.51515	5.3
Aomori	1,321	27	23	4	0.302800908	28993.18698	1.1
Kagoshima	1,668	11	6	5	0.299760192	28657.07434	4.8
Gifu	2,041	31	25	6	0.293973542	27290.54385	2.7
Kumamoto	1,794	19	14	5	0.2787068	28093.64548	3.5
Kochi	738	12	10	2	0.27100271	32113.82114	2.5
Tokushima	764	12	10	2	0.261780105	30104.71204	2.8
Hiroshima	2,833	43	36	7	0.247087893	27144.36993	5.3
Kyoto	2,610	55	49	6	0.229885057	26858.23755	5.1
Aichi	7,455	162	145	17	0.228034876	23179.07445	3.7
Osaka	8,836	150	130	20	0.226346763	25656.40561	3.1
Hyōgo	5,541	82	70	12	0.216567407	26349.03447	4.7
Chiba	6,197	73	60	13	0.209778925	25350.97628	2.9
Wakayama	971	4	2	2	0.205973223	30484.03708	3.3
Ibaraki	2,919	41	35	6	0.205549846	25830.76396	1.8
Oita	1,171	22	20	2	0.170794193	29632.79249	2.4
Fukuoka	5,091	40	32	8	0.157140051	25122.76566	2.5
Okayama	1,924	36	33	3	0.155925156	28066.52807	6
Kanagawa	9,096	63	49	14	0.153913808	23251.97889	5.7
Yamaguch	1,408	17	15	2	0.142045455	31321.02273	1.9
Shiga	1,416	30	28	2	0.141242938	23446.32768	1.7
Yamanashi	841	5	4	1	0.118906064	27467.30083	3.9
Triple	1,825	20	18	2	0.109589041	27123.28767	0.8
Gunma	1,976	26	24	2	0.101214575	26771.25506	1.4
Akita	1,037	8	7	1	0.096432015	32690.45323	2.4
Miyazaki	1,114	5	4	1	0.089766607	28635.54758	1.4
Ishikawa	1,156	28	27	1	0.08650519	27076.12457	2.3
Saitama	7,239	37	31	6	0.082884376	23995.02694	4
Nara	1,376	16	15	1	0.072674419	27834.30233	2.7
Okinawa	1,421	16	15	1	0.070372977	19000.70373	3.5