



# Regional food purchasing behavior characteristics contained in “Family Income and Expenditure Survey” in Japan

uRos2024 Greece the 29<sup>th</sup> of November 2024 presentation

Atsushi Kimura

e-mail: [atkimura@nstac.go.jp](mailto:atkimura@nstac.go.jp)

GitHub: [https://github.com/ibuchichi/R\\_function\\_2024.git](https://github.com/ibuchichi/R_function_2024.git)

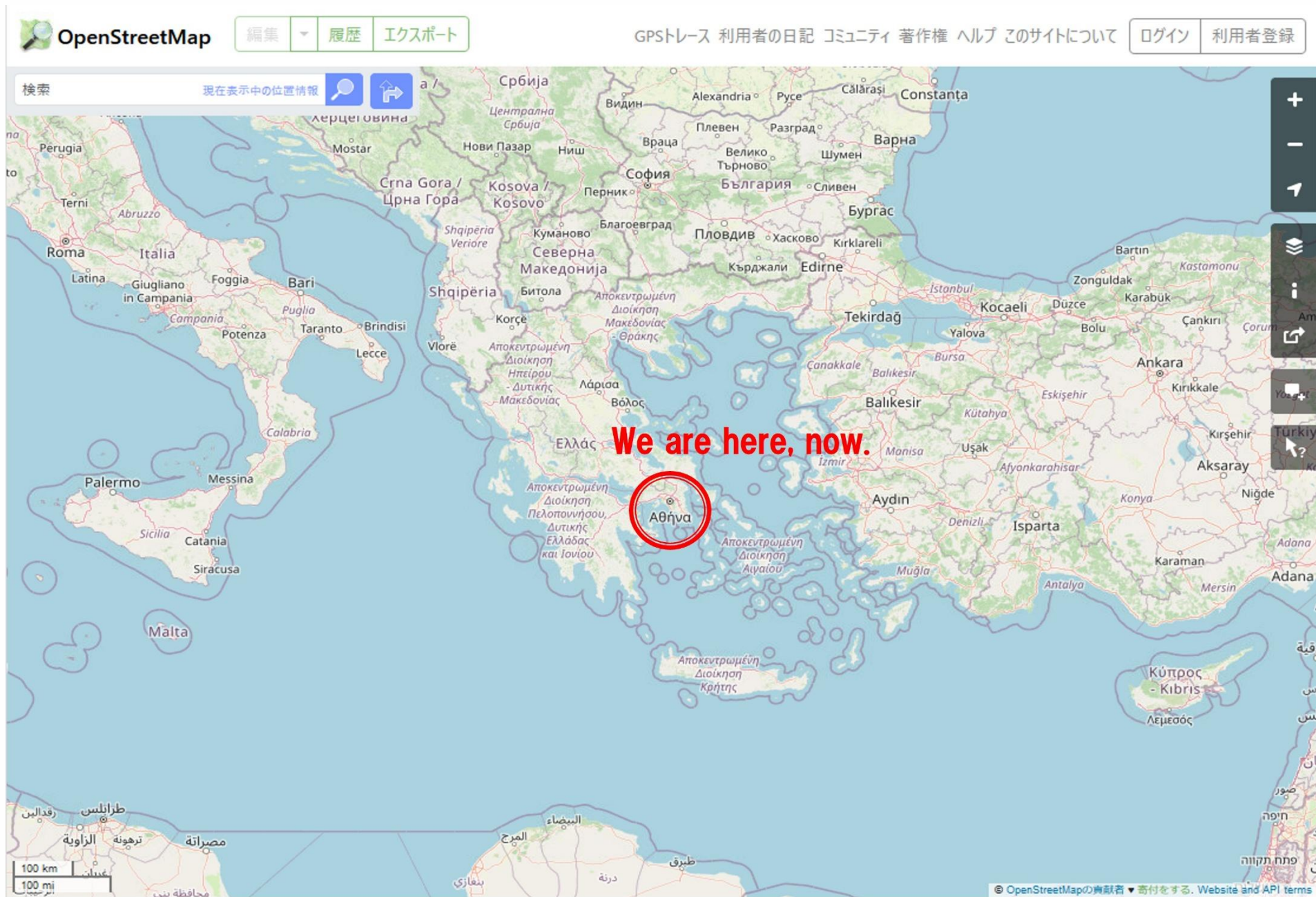
National Statistics Center, Japan

<https://www.nstac.go.jp>

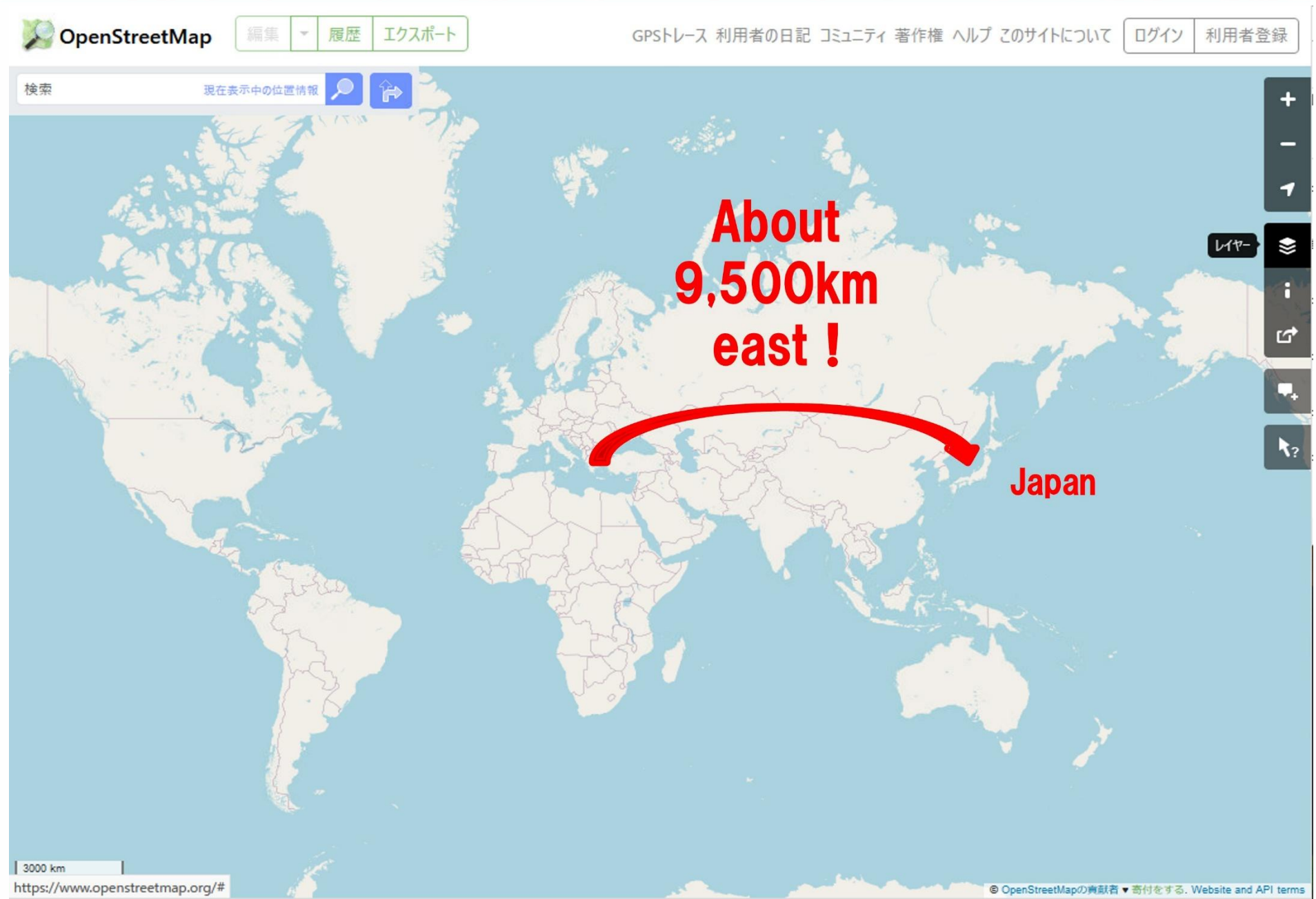
- 1. Analysis of Japan's "Family Income and Expenditure Survey" reveals regional characteristics of food purchasing behavior**
- 2. The possibility of "GSBPM 6.2 Validate outputs" using these regional characteristics**
- 3. Two useful tools for comparing and evaluating the results of multiple geographic cluster analyses**

- 1. Analysis of Japan's "Family Income and Expenditure Survey" reveals regional characteristics of food purchasing behavior**
2. The possibility of "GSBPM 6.2 Validate outputs" using these regional characteristics
3. Two useful tools for comparing and evaluating the results of multiple geographic cluster analyses

# About Japan



# About Japan



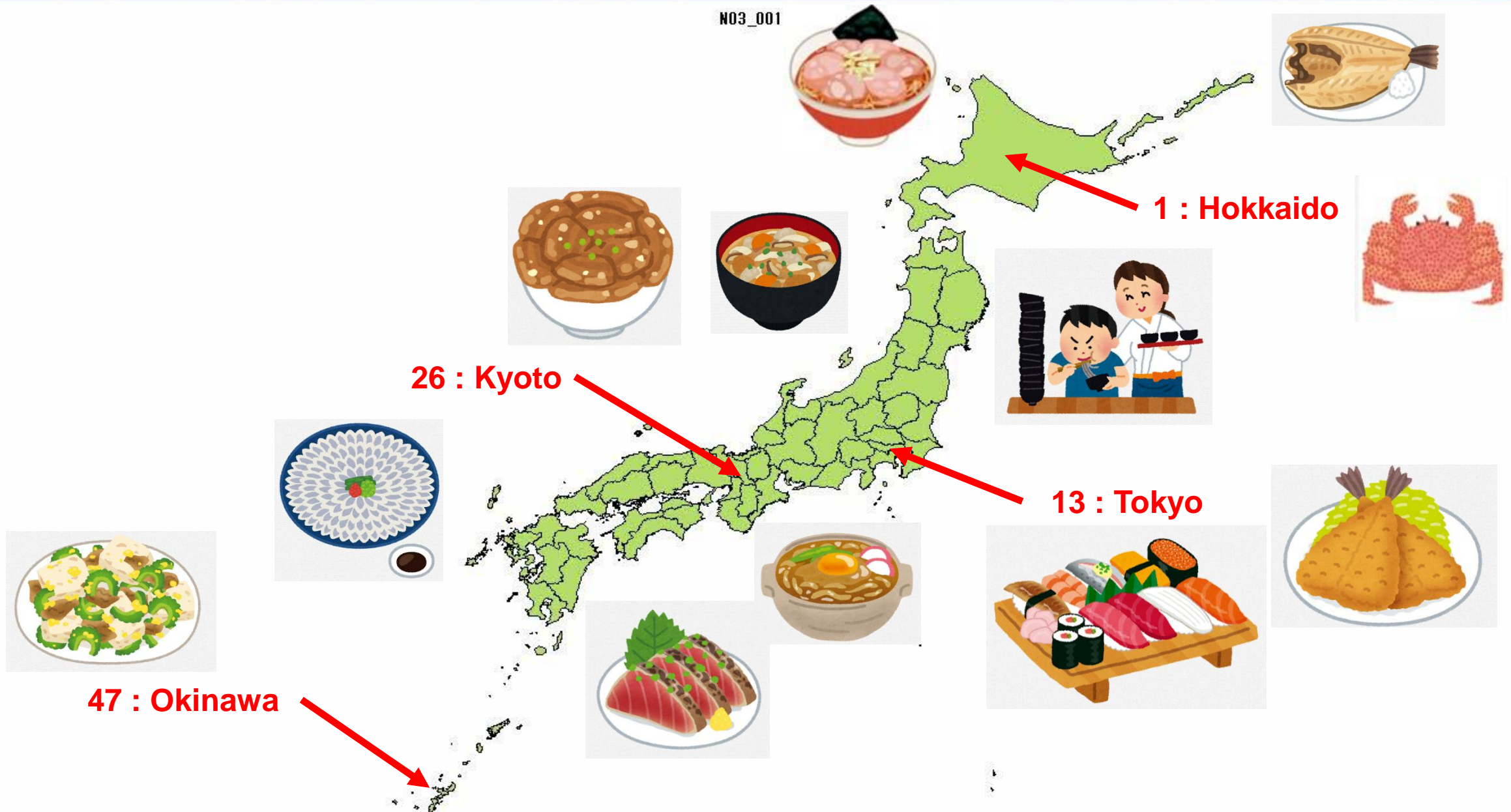
The image is a screenshot of the OpenStreetMap website. At the top left, the OpenStreetMap logo is visible, along with buttons for '編集' (Edit), '履歴' (History), and 'エクスポート' (Export). To the right, there are links for 'GPSトレース' (GPS Trace), '利用者の日記' (User Diaries), 'コミュニティ' (Community), '著作権' (Copyright), 'ヘルプ' (Help), and 'このサイトについて' (About this site). Further right are buttons for 'ログイン' (Login) and '利用者登録' (User Registration). Below the navigation bar is a search bar with the text '検索' (Search) and '現在表示中の位置情報' (Current location information). The main area is a world map. A red curved arrow points from the western coast of Europe to Japan. Above the arrow, the text 'About 9,500km east!' is written in red. To the right of the arrow, the word 'Japan' is written in red. On the right side of the map, there is a vertical toolbar with icons for zooming in (+), zooming out (-), home, layers (レイヤー), and other map controls. At the bottom left, there is a scale bar for 3000 km and the URL 'https://www.openstreetmap.org/#'. At the bottom right, there is a copyright notice: '© OpenStreetMapの貢献者 ▼寄付をする. Website and API terms'.

# About Japan



# 47 Prefectures

N03\_001



- The survey covers about 9,000 households in 168 cities, towns, and villages in Japan (randomly selected from all over the country using statistical methods).
- Conducted by the national government through prefectural governments (survey conducted by surveyors; fundamental statistical survey).



## Method of investigation

- Survey households fill out a "household account book" regarding their daily income and expenditures. (Online surveys are also conducted via PC, smartphone, etc.)

## Matters for investigation

- Consumption expenditure by item (e.g., town revitalization....No. 1 consumption of gyoza in Japan, etc.)
- Breakdown of actual income of working households
- Savings, debt, etc.
- The total number of income and expense items is approximately 800.



<https://www.stat.go.jp/english/data/kakei/index.html>

# Data Table (2020 to 2022)

## 212 food items

	Rice	White bread	Other bread	Non-dried "Udon"& "Soba"	Dried "Udon"& "Soba"	Pasta	Chinese noodles	...	Tuna fish	Horse mackerel	Sardines	Bonito	Flounder	Salmon	...	Beef	Pork	Chicken	...	Yogurt	Butter	Cheese	...	Cabbage	Spinach	Chinese cabbage	Welsh onions	...	School lunch	
HOKKAIDO	27317	9192	18056	3218	2332	1290	4285	...	5100	186	203	1046	1441	7877	...	13779	3060	17045	...	11565	1743	8023	...	3031	2479	1515	3375	...	12578	
AOMORI	20976	8441	17808	3548	2177	1229	5851	...	4620	634	563	1294	1622	6347	...	14945	3798	16168	...	12730	1322	5731	...	3354	2430	1231	3214	...	10485	
IWATE	22138	8585	20439	3360	2309	1358	6540	...	5659	464	344	2267	1110	5557	...	11381	3485	14474	...	15712	1185	7155	...	3174	3092	1191	3007	...	10519	
MIYAGI	19328	9344	21631	3202	1990	1395	5181	...	6203	550	270	2350	988	5130	...	13993	3479	15406	...	15426	1322	7565	...	3010	2673	1317	3044	...	10094	
AKITA	19497	7874	16883	3532	4397	1167	5573	...	4984	933	516	1424	2232	5868	...	14839	3038	15362	...	13533	1072	6394	...	3347	3045	1226	3188	...	11509	
YAMAGATA	25373	8182	17530	4549	3306	1338	5948	...	5803	431	191	2335	804	4296	...	27068	3242	16057	...	15852	1178	6568	...	2686	2494	914	3245	...	12801	
FUKUSHIMA	22051	8745	18299	3303	2490	1330	5712	...	5472	472	37	13	4796	...	13105	3895	13865	...	17005	997	6748	...	3440	2274	1304	2980	...	13226		
IBARAKI	19319	9375	17860	3626	2257	1397	4256	...	5450	577	285	2628	799	5175	...	14480	3336	14335	...	14427	1350	7126	...	2486	1799	998	2899	...	8907	
TOCHIGI	19441	10835	19014	3735	2173	1037	4331	...	7324	348	202	1348	538	4833	...	13482	3045	14311	...	13243	1278	7354	...	2328	1538	937	2338	...	1752	
GUNMA	23138	9484	22827	4737	2441	1381	4726	...	7661	658	356	1232	624	5836	...	13016	2698	12899	...	16198	1421	7092	...	2659	1775	909	2639	...	12479	
SAITAMA	21150	10519	23277	3897	2100	1694	5111	...	7564	906	344	1300	591	5739	...	19665	3639	17725	...	14587	1602	8608	...	2936	2225	1115	3667	...	12916	
CHIBA	21175	11661	23472	3640	2226	1576	4746	...	7609	1154	472	1623	791	6116	...	20973	3668	16053	...	15601	1662	7904	...	2972	2427	1199	4066	...	10414	
TOKYO	20664	11722	23149	3298	2273	1786	5058	...	8123	1186	400	1571	630	5759	...	26325	3972	18529	...	15138	1861	9367	...	3058	2444	1286	4204	...	8111	
KANAGAWA	23136	11782	22607	3638	2998	1747	5122	...	8005	1159	481	1353	598	5374	...	23538	3485	18297	...	14833	1624	8454	...	3409	2412	1302	4267	...	6589	
NIGATA	25355	10760	21955	2809	3174	1627	5014	...	3639	705	341	1096	1404	6398	...	11607	3309	14585	...	14860	1230	7367	...	3291	2398	1288	3379	...	13044	
TOYAMA	26113	10814	21980	3472	3105	1253	5019	...	4795	1555	426	1120	840	6184	...	18938	3982	14520	...	14652	1303	6716	...	3178	2486	1408	3089	...	11188	
ISHIKAWA	26087	11708	22226	3633	2260	1466	5050	...	3192	1489	885	933	1576	4950	...	24452	3903	17641	...	11685	1487	7912	...	2714	2373	1267	3350	...	18544	
FUKUI	26910	9114	21715	4115	2724	1256	4509	...	2729	818	485	787	1454	4732	...	24526	2914	16351	...	12073	1064	5438	...	2552	2636	1148	2920	...	11631	
YAMANASHI	22365	8948	21230	4017	1812	1296	5003	...	8885	745	168	993	452	4773	...	16006	3492	14238	...	14999	1195	6231	...	2585	1913	1126	3009	...	10304	
NAGANO	20489	8121	20055	4322	2369	1370	4614	...	5164	575	279	907	995	5380	...	12641	3539	12634	...	13025	1234	7336	...	2751	1719	1039	2068	...	12685	
GIFU	22443	10952	20004	3378	2099	1348	4193	...	4241	639	381	993	540	5221	...	23401	2534	14871	...	14920	1268	6677	...	2430	2148	1053	3183	...	12522	
SHIZUOKA	28588	10017	22429	3079	2283	1413	5232	...	10974	1119	255	1555	491	4938	...	17469	3551	16665	...	14669	1523	6876	...	3576	2208	1369	4306	...	10979	
AICHI	21965	11622	21246	4264	2092	1465	4863	...	5481	847	367	929	516	4842	...	24176	3596	16794	...	14474	1471	6361	...	2830	2028	1389	3882	...	10310	
MIE	18032	10146	21954	3736	2121	1397	4597	...	5983	1061	556	1537	456	4911	...	32495	2167	14712	...	14712	1471	7435	...	2630	1683	1165	3561	...	9281	
SHIGA	22595	11686	25294	4107	2151	1343	4420	...	3973	999	443	1401	1429	5280	...	36702	3356	20360	...	14285	1494	7435	...	2813	2620	1529	4023	...	12764	
KYOTO	23483	12693	26706	4227	1929	1427	4441	...	3861	1073	457	1511	1567	5482	...	39377	3010	20995	...	13091	1615	7334	...	3117	2485	1738	4077	...	8643	
OSAKA	23418	11116	23280	3857	1518	1122	3919	...	4382	914	332	1522	1356	4818	...	35113	3755	18831	...	12272	1270	6008	...	2773	2234	1774	3903	...	2205	
:	:	:	:	:	:	:	:	...	:	:	:	:	:	:	...	:	:	:	...	:	:	:	...	:	:	:	:	:	...	:
KAGOSHIMA	22570	8061	19697	2823	2280	1308	3587	...	2840	1486	859	1430	395	5123	...	23980	3729	20850	...	12459	1001	5101	...	3209	1549	1116	2871	...	14191	
OKINAWA	27578	8788	18726	1934	1592	1065	4951	...	5159	205	134	732	138	4809	...	14508	28974	13172	...	10438	941	6750	...	2920	2013	1105	2598	...	9919	

Convert to Ratio

Standardize

47 prefectures

# Preprocessed Data Table (2020 to 2022)



212 food items

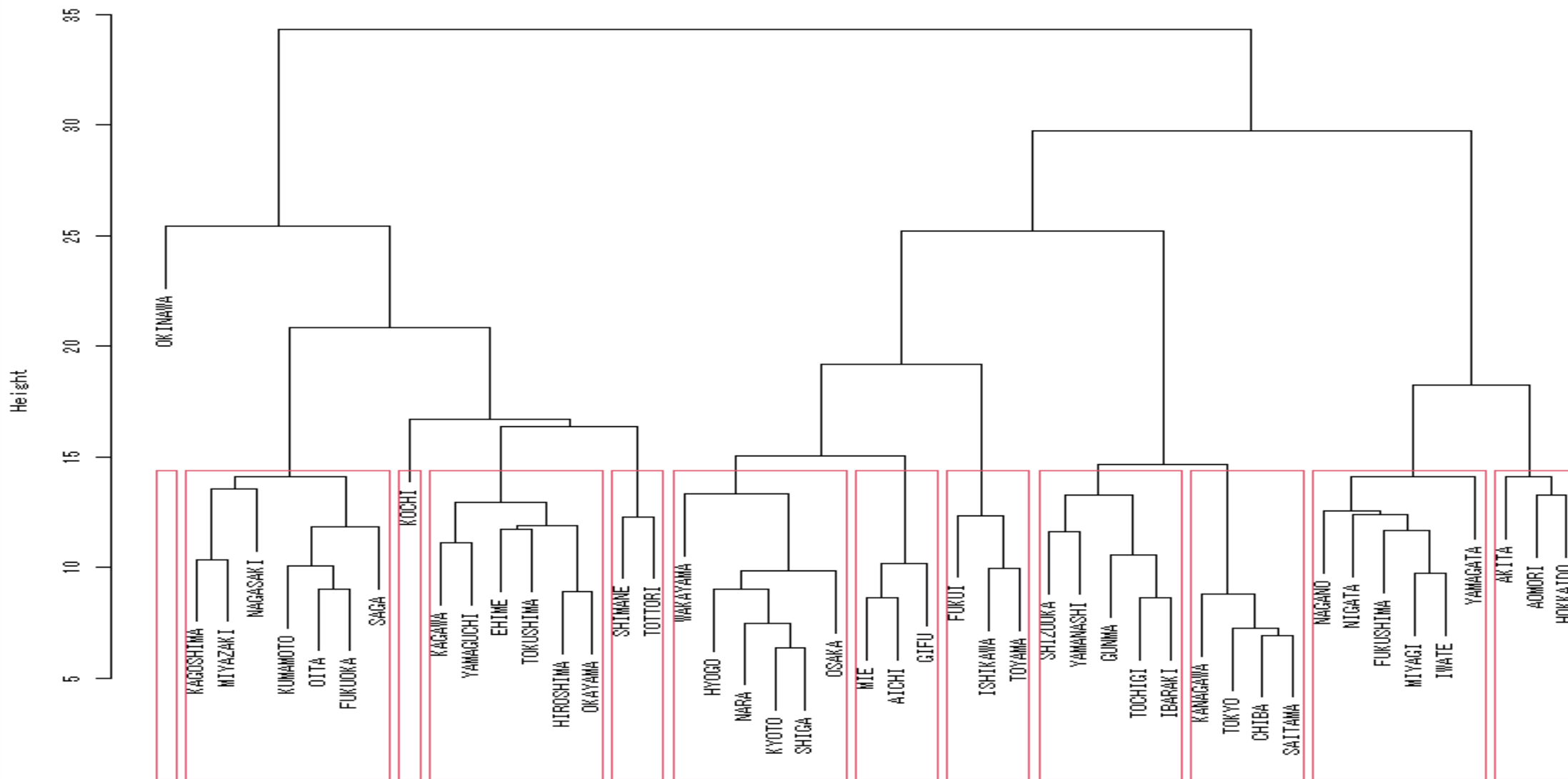
47 prefectures

	Rice	White bread	Other bread	Non-dried "Udon"& "Soba"	Dried "Udon"& "Soba"	Pasta	Chinese noodles	...	Tuna fish	Horse mackerel	Sardines	Bonito	Flounder	Salmon	...	Beef	Pork	Chicken	...	Yogurt	Butter	Cheese	...	Cabbage	Spinach	Chinese cabbage	Welsh onions	...	School lunch	
HOKKAIDO	1.957	-0.694	-1.279	-0.3753238	0.2488197	-0.245	-0.16763	...	0.369	-1.583925	-1.173689	-0.4669	0.8636264	4.094663	...	-1.219807	0.999	0.02351	...	-1.3756	2.76775	1.985165	...	0.3454436	1.169757	1.328651	0.5674	...	0.43092	
AOMORI	-0.033	-1.179	-1.195	0.22495758	0.0597521	-0.479	2.340357	...	0.195	-0.765918	0.7001327	-0.1976	1.2619641	2.114639	...	-1.017984	1.279	-0.16227	...	-0.1947	0.5297	-0.5653	...	1.5607872	1.196569	0.006738	0.3458	...	-0.1473	
IWATE	0.16	-1.264	-0.24	-0.1895193	0.1894857	0.2293	3.104609	...	0.624	-1.096456	-0.473465	0.70073	0.2581882	0.722787	...	-1.541751	0.117	-0.98871	...	1.89382	-0.4714	0.923566	...	0.7377005	2.728033	-0.38203	-0.452	...	-0.23	
MIYAGI	-0.978	-0.821	0.0252	-0.5438538	-0.502631	0.2163	0.907576	...	0.8	-0.972106	-0.880466	0.71777	-0.0109	0.799999	...	0.058878	1.174	-0.56082	...	1.32452	0.10381	1.15965	...	-0.002492	1.454209	0.081741	-0.579	...	-0.4511	
AKITA	-0.641	-1.792	-1.726	0.1321989	4.1710527	-1.078	1.811263	...	0.34	-0.247696	0.4235885	-0.088	2.335532	1.293214	...	0.058878	1.174	-0.56082	...	0.31637	-1.0287	0.136667	...	1.4027604	2.722801	-0.10664	0.1579	...	0.13369	
YAMAGATA	0.824	-1.988	-1.945	1.187052	1.7244	-0.43	1.799824	...	0.554	-1.192755	-1.285761	0.65278	-0.36525	-1.39736	...	0.3386263	0.649	-0.69252	...	1.36706	-0.8529	-0.16115	...	-1.127119	0.858123	-2.05785	-0.254	...	0.2827	
FUKUSHIMA	-0.167	-1.43	-1.527	-0.4590073	0.3418994	-0.313	0.078037	...	0.903	-1.115246	-0.397592	1.42924	0.0115845	-0.64119	...	-1.383497	1.664	-1.43313	...	2.38777	-1.776	0.127683	...	1.1257942	0.382825	-0.06379	-0.84	...	0.45455	
IBARAKI	-0.501	-0.187	-1.036	0.42792912	0.2823661	0.9988	0.108499	...	0.642	-0.85185	-0.687993	1.16132	-0.236094	0.511042	...	-1.048682	0.234	-0.79288	...	1.39042	0.82555	1.270346	...	-0.982707	-0.401	-1.14843	-0.39	...	-0.6089	
TOCHIGI	-1.099	0.2587	-0.886	0.18123866	-0.18755	2.0895	0.555531	...	1.395	-0.811962	-0.827091	-0.0356	-0.710239	-0.53227	...	-0.93951	-0.111	-0.99974	...	1.14515	-0.167	0.887322	...	-0.277054	-0.33872	-1.54949	-0.903	...	0.04149	
GUNMA	0.27	-0.713	0.5328	1.58824135	0.301815	0.099	0.250979	...	1.465	-0.789162	-0.465438	-0.3269	-0.648705	0.878607	...	-1.376946	-1.257	-1.73771	...	1.90428	0.64416	0.604777	...	-1.034191	-0.80726	-1.98193	-1.643	...	0.2703	
SAITAMA	-1.035	-0.662	-0.213	-0.0861687	-0.660425	1.2365	0.123976	...	1.098	-0.509443	-0.678453	-0.376	-0.802129	0.001596	...	-0.762049	0.522	-0.55269	...	-0.3901	0.82387	1.426357	...	-1.021325	-0.19687	-1.466	0.1428	...	0.03834	
CHIBA	-0.942	0.3513	-0.012	-0.3493461	-0.406627	0.5946	-0.26408	...	1.16	-0.101522	-0.08766	-0.0846	-0.468537	0.581383	...	-0.578279	-1.089	-1.04512	...	0.4684	1.23941	0.804544	...	-0.821663	0.336563	-1.00902	1.2114	...	-0.614	
TOKYO	-1.507	-0.207	-0.748	-1.0637456	-0.574477	1.2202	-0.28871	...	1.152	-0.173049	-0.525209	-0.2172	-0.793137	-0.35795	...	-0.152867	-0.65	-0.61221	...	-0.5654	1.63486	1.700352	...	-1.125994	0.007215	-0.99573	0.8842	...	-1.4069	
KANAGAWA	-0.433	0.3323	-0.47	-0.4092694	0.8066625	1.5984	0.144242	...	1.285	-0.116063	-0.073936	-0.3301	-0.790164	-0.45321	...	-0.311625	0.924	-0.35363	...	-0.2078	0.94258	1.274382	...	0.2324645	0.233547	-0.60652	1.5659	...	-1.707	
NIGATA	0.844	0.2109	-0.028	-1.1717991	1.510347	1.6717	0.525511	...	-0.413	-0.731638	-0.569345	-0.4728	0.6670252	1.493184	...	-1.584477	1.721	-1.21201	...	0.65211	-0.5484	0.760708	...	0.6201029	0.641413	-0.19631	0.1135	...	0.36739	
TOYAMA	0.756	-0.104	-0.388	-0.4526526	1.1731089	-1.363	0.256862	...	0.023	0.5890445	-0.244922	-0.4913	-0.355959	0.867601	...	-0.76094	-1.034	-1.4451	...	0.05553	-0.4304	-0.26913	...	-0.058623	0.616721	0.122139	-0.936	...	-0.3186	
ISHIKAWA	0.698	0.5633	-0.343	-0.2705507	-0.280159	0.0416	0.256832	...	-0.68	0.4681693	1.8275495	-0.6604	0.8392845	-0.79181	...	-0.116502	0.248	-0.38722	...	-2.1605	0.48721	0.970191	...	-1.373535	0.315251	-0.58374	-0.347	...	1.75884	
FUKUI	1.821	-0.762	0.3733	0.91863261	0.9779538	-0.499	0.162896	...	-0.756	-0.46811	0.2407977	-0.7162	0.8878062	-0.43234	...	0.1978183	-1.671	-0.24219	...	-0.9665	-1.14	-1.07667	...	-1.090238	1.578665	-0.58021	-0.653	...	0.13383	
YAMANASHI	0.16	-1.022	0.0187	0.69530068	-0.763409	-0.335	0.783938	...	2.108	-0.615635	-1.360887	-0.5313	-0.933494	-0.47078	...	-0.95559	0.469	-1.13434	...	1.20507	-0.4831	-0.24136	...	-1.101226	-0.36679	-0.77791	-0.528	...	-0.33	
NAGANO	-0.227	-1.496	-0.187	1.34637253	0.4101096	0.5776	0.45778	...	0.451	-0.876007	-0.763207	-0.5824	0.0961425	0.66206	...	-1.334743	0.023	-1.57353	...	0.01201	-0.0133	1.354019	...	-0.321147	-0.70693	-1.03536	-2.826	...	0.54866	
GIFU	-0.057	0.4263	-0.813	-0.367399	-0.361993	-0.272	-0.58411	...	-0.131	-0.836	-0.368636	-0.5617	-0.807836	-0.07387	...	-0.092453	-1.535	-1.07617	...	0.76286	-0.301	0.02805	...	-1.792084	0.056819	-1.32706	-0.339	...	0.23303	
SHIZUOKA	1.789	-0.51	0.0689	-0.849164	-0.103543	0.0263	0.748664	...	2.837	-0.059054	-0.991826	-0.069	-0.906391	-0.57605	...	-0.875382	1.154	-0.51888	...	0.3842	0.96032	0.123266	...	1.3193496	0.114056	0.129135	2.3548	...	-0.2855	
AICHI	-0.225	0.9789	-0.294	0.84090961	-0.380603	0.5535	0.342818	...	0.427	-0.488077	-0.438557	-0.622	-0.850678	-0.60582	...	-5.32E-05	-1.034	-0.37949	...	0.4038	0.8008	-0.34137	...	-0.659954	-0.24849	0.330476	1.4357	...	-0.4366	
MIE	-1.133	0.272	0.6069	0.44084673	-0.09181	0.7006	0.377825	...	0.825	-0.015321	0.6322856	0.02524	-0.904228	-0.08463	...	1.30308	-0.827	0.01819	...	1.04396	-0.6247	-0.66222	...	-0.755728	-0.83747	-0.41461	1.1912	...	-0.5689	
SHIGA	-0.37	0.5609	0.9268	0.35677987	-0.457552	-0.785	-0.57106	...	-0.341	-0.312554	-0.176394	-0.2501	0.6026808	-0.35014	...	1.3534007	-0.926	0.57329	...	-0.2592	0.53644	0.471174	...	-1.09189	0.905381	0.663414	1.3059	...	0.1169	
KYOTO	-0.049	1.4398	1.5744	0.54899606	-0.810695	-0.147	-0.50526	...	-0.379	-0.183067	-0.09941	-0.1455	0.8451768	-0.0404	...	1.7034764	-0.993	0.84222	...	-1.0652	1.22248	0.413039	...	-0.209717	0.625222	1.703026	1.5011	...	-1.0441	
OSAKA	0.234	0.5329	0.5615	0.27278359	-1.400641	-1.929	-0.98686	...	-0.074	-0.378181	-0.60857	-0.0803	0.5900679	-0.6506	...	1.3655919	-0.489	0.3558	...	-1.2915	-0.3142	-0.75209	...	-0.837164	0.250446	2.224	1.4715	...	-2.8681	
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
KAGOSHIMA	0.654	-1.406	-0.174	-0.7771243	0.3263934	0.2978	-0.97549	...	-0.649	0.8183433	2.2923729	-0.0404	-0.995719	0.432317	...	0.2550013	0.986	1.81997	...	-0.2502	-1.2673	-1.22913	...	1.2759132	-1.07634	-0.50944	-0.469	...	1.12235	
OKINAWA	3.129	-0.155	-0.007	-1.9050064	-0.785336	-1.019	1.644925	...	0.679	-1.508808	-1.442413	-0.6884	-1.460463	0.472617	...	-0.903888	0.606	-0.88543	...	-1.3175	-1.2271	1.395911	...	1.0196025	0.561383	-0.14454	-0.718	...	-0.0472	

std\_data ←

```
std_data.dist <- sqrt(1/2)*dist(std_data)  
hclust.result <- hclust(std_data.dist, method="ward.D2")
```

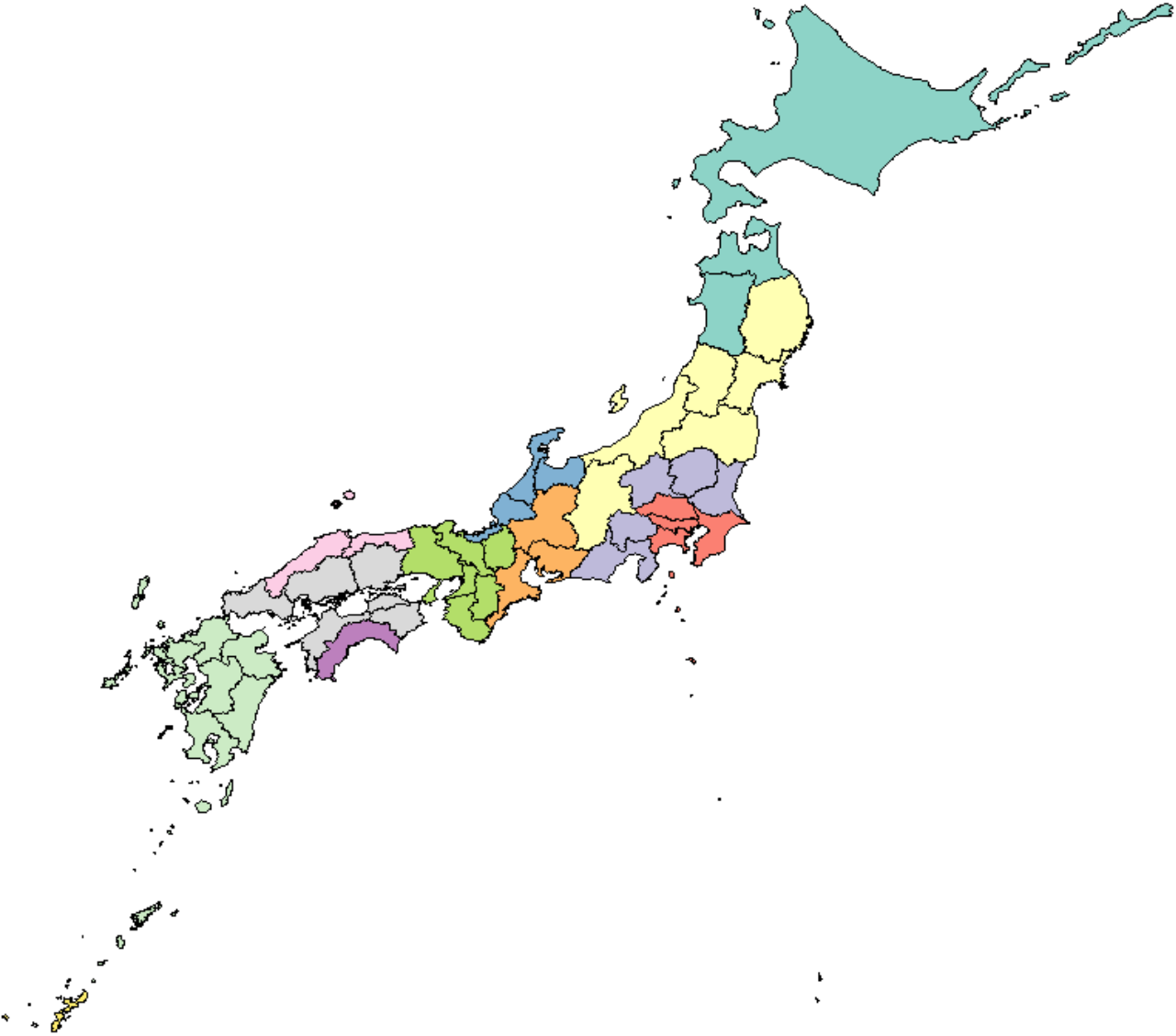
Cluster Dendrogram



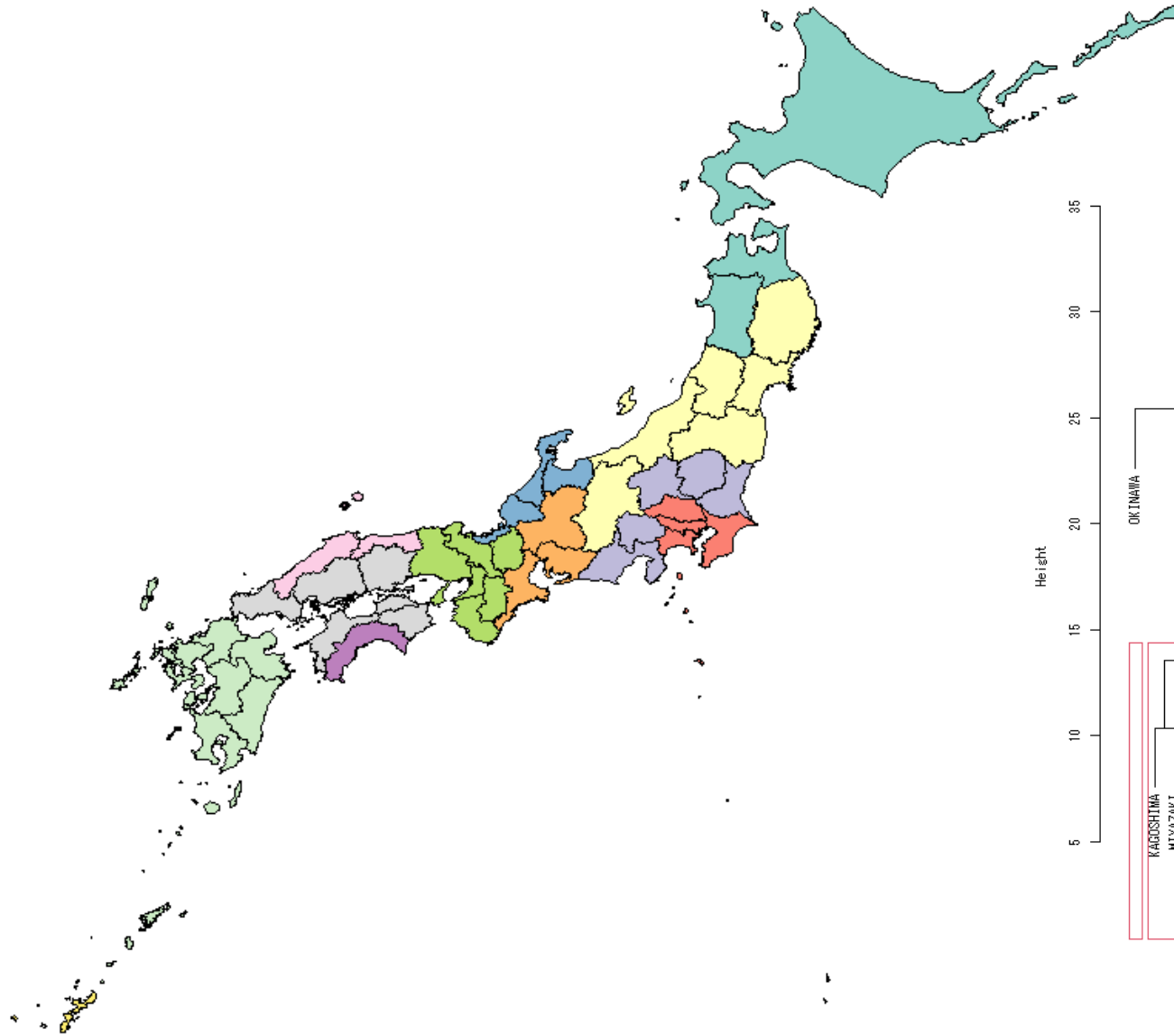
ssdse\_small.dist  
hclust (\*, "ward.D2")

# 12 clusters Map (2020 to 2022)

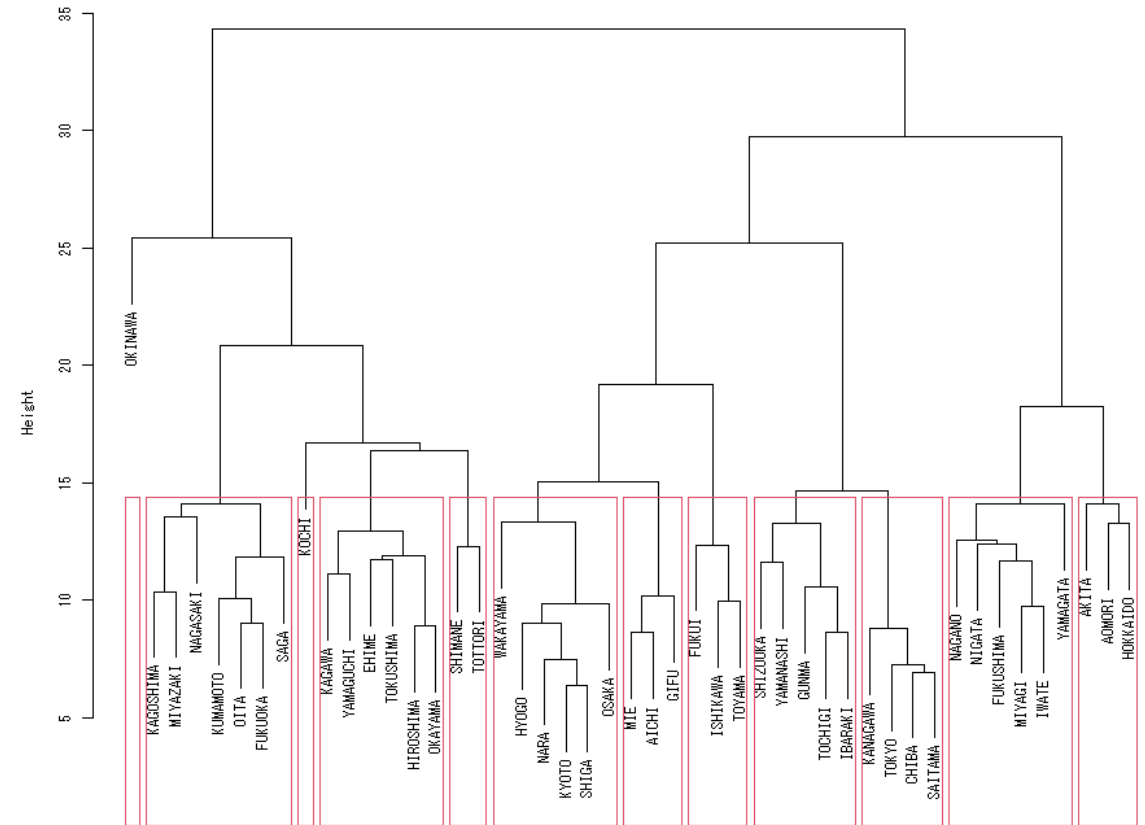
N03\_001



N03\_001

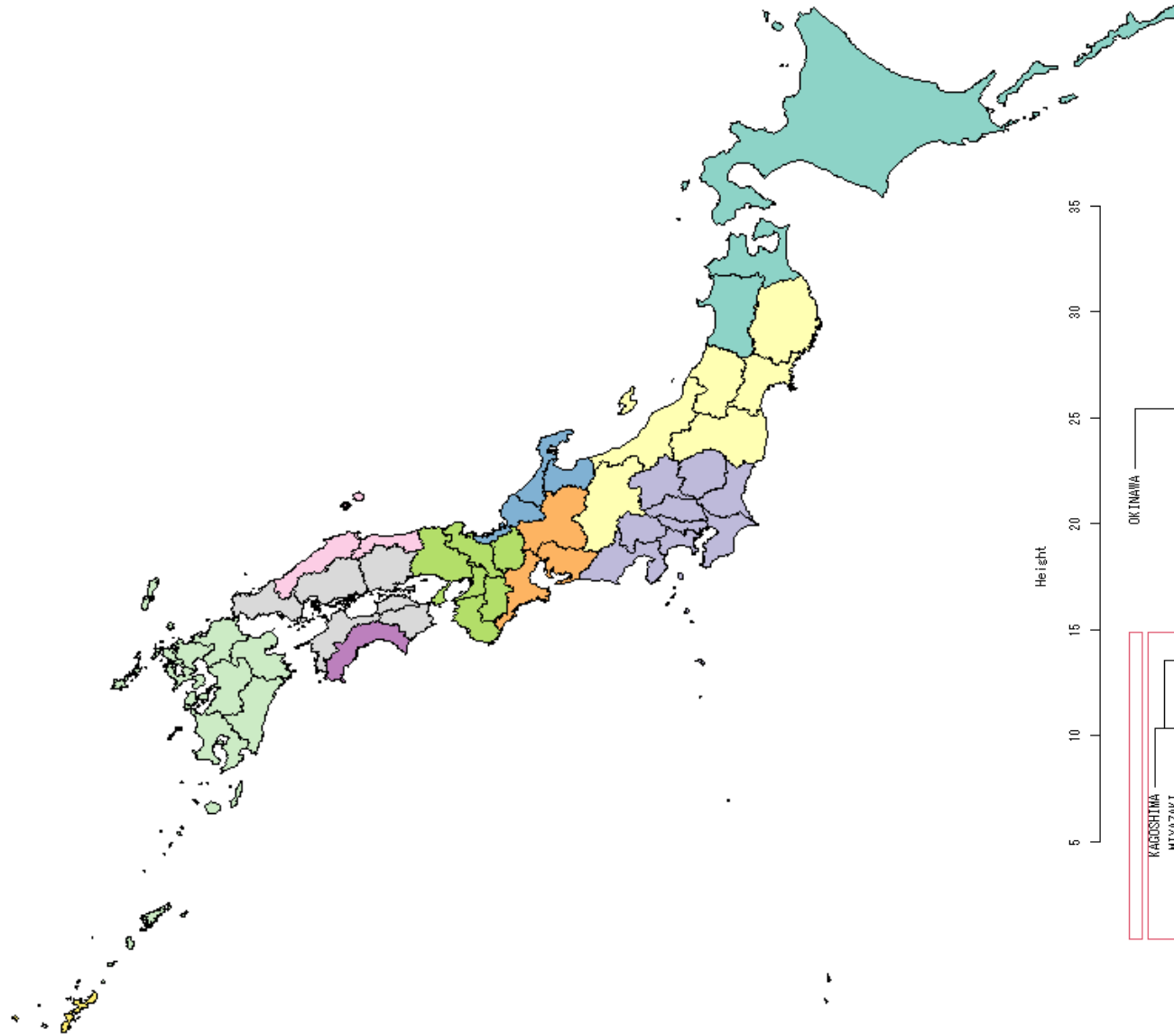


Cluster Dendrogram

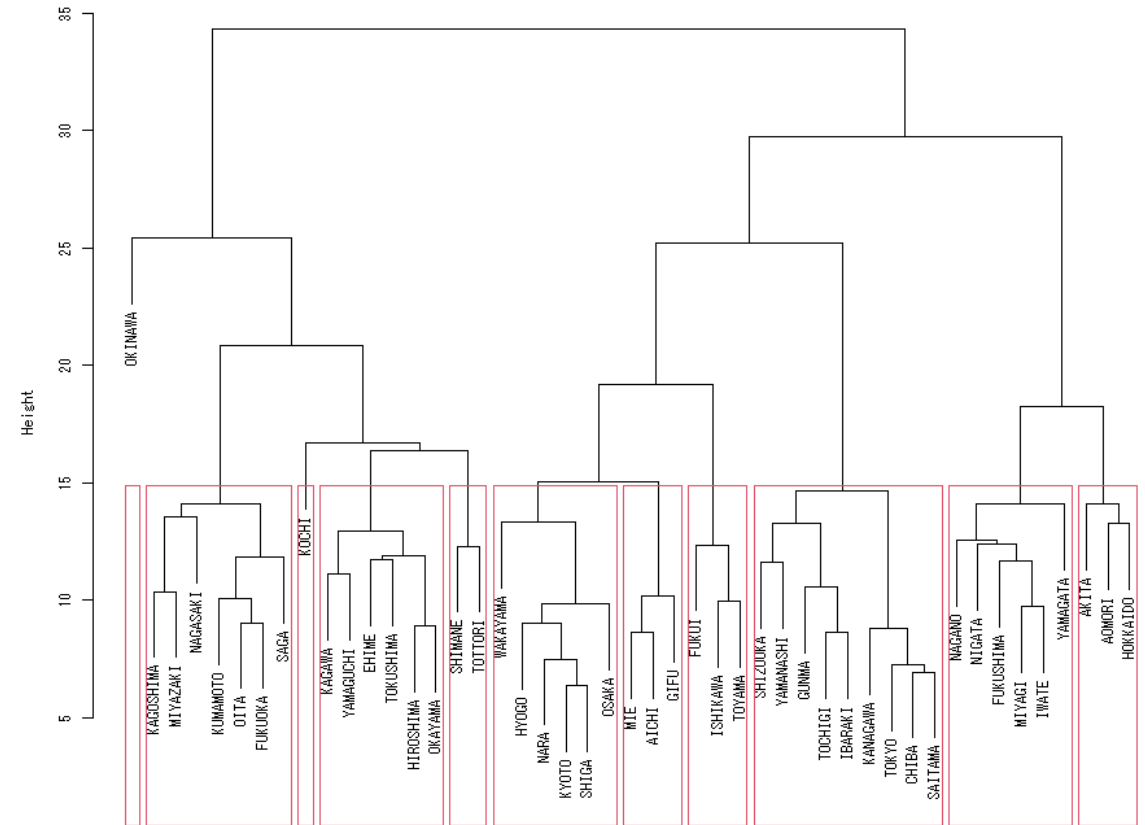


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

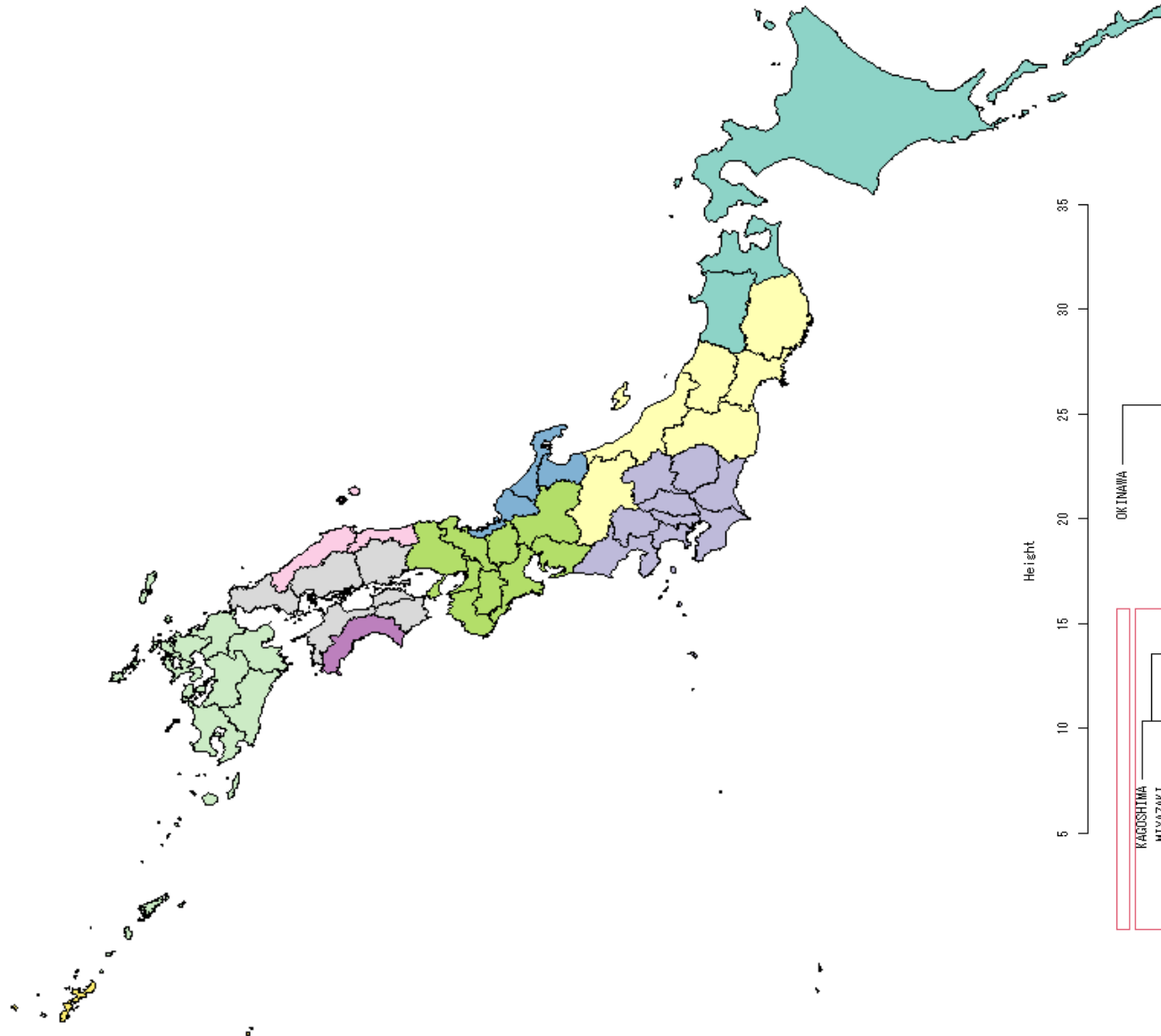


Cluster Dendrogram

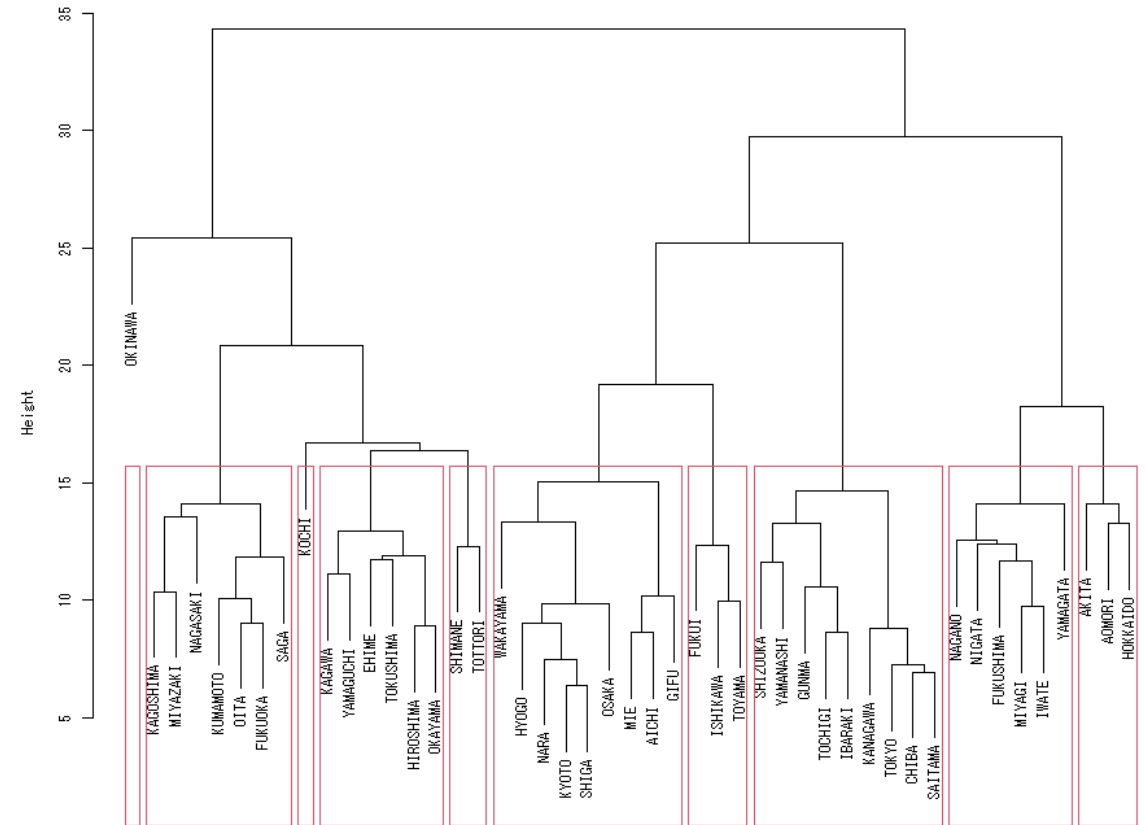


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

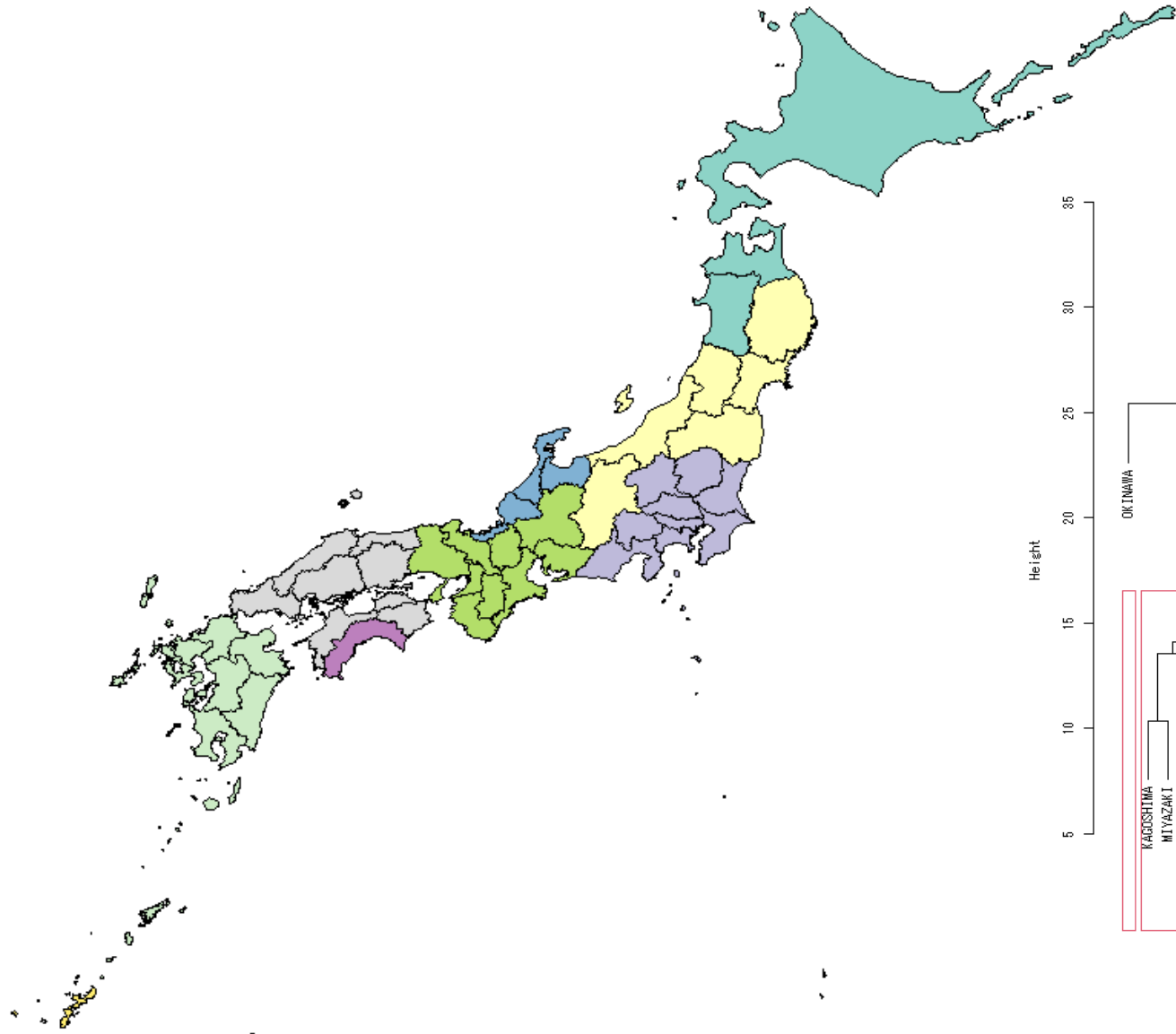


Cluster Dendrogram

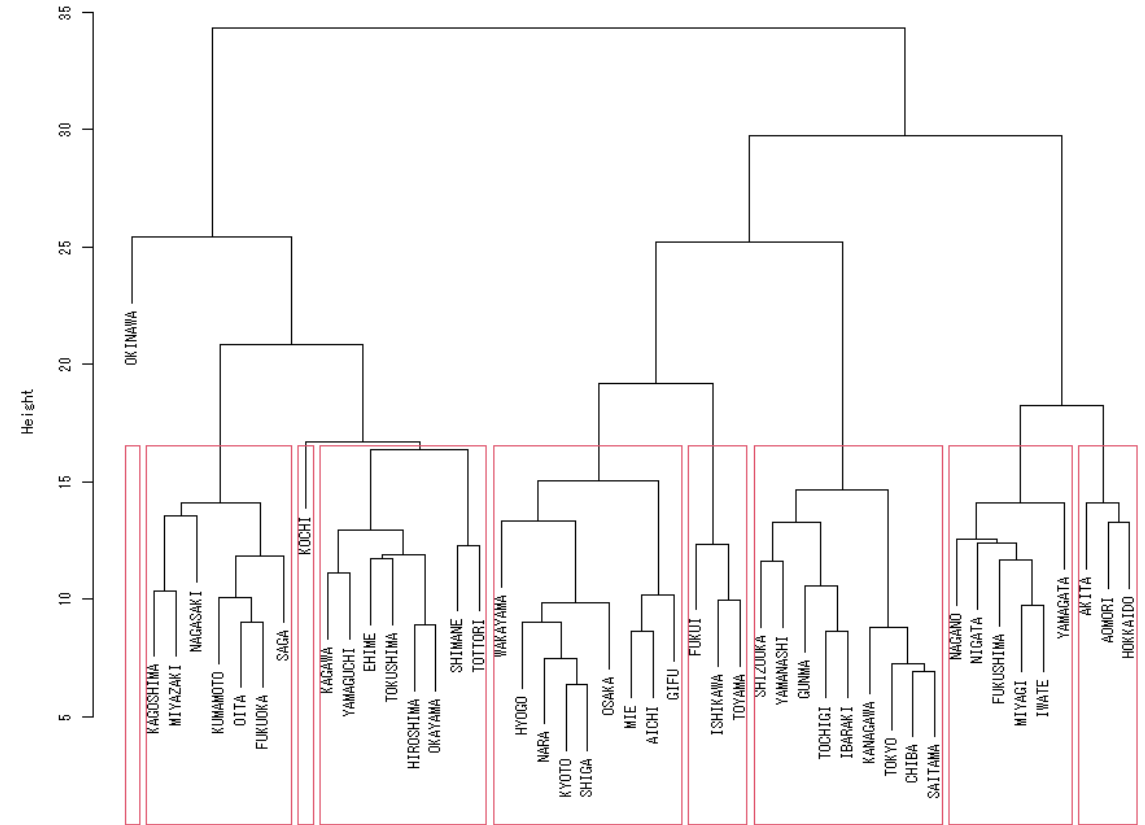


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

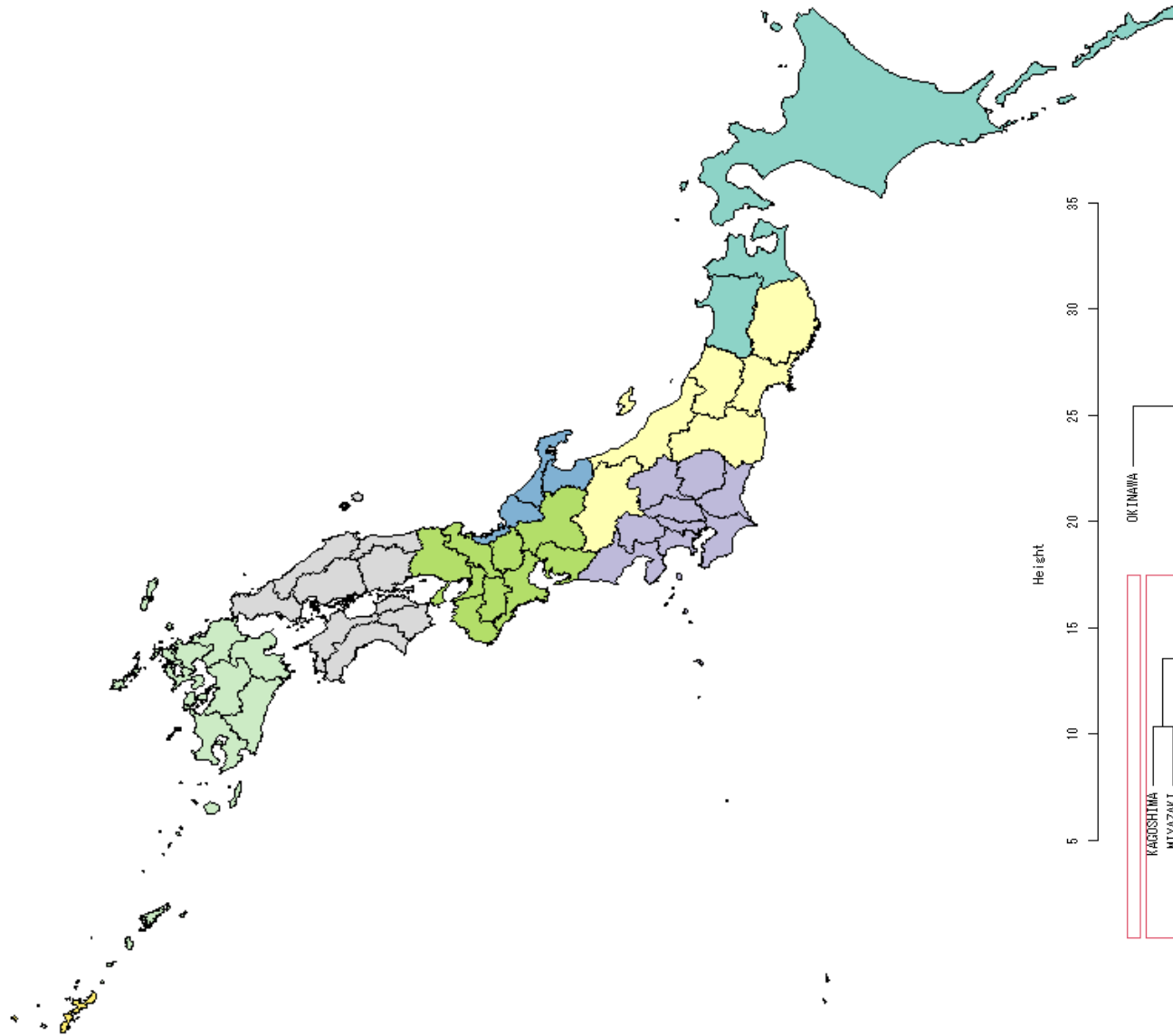


Cluster Dendrogram

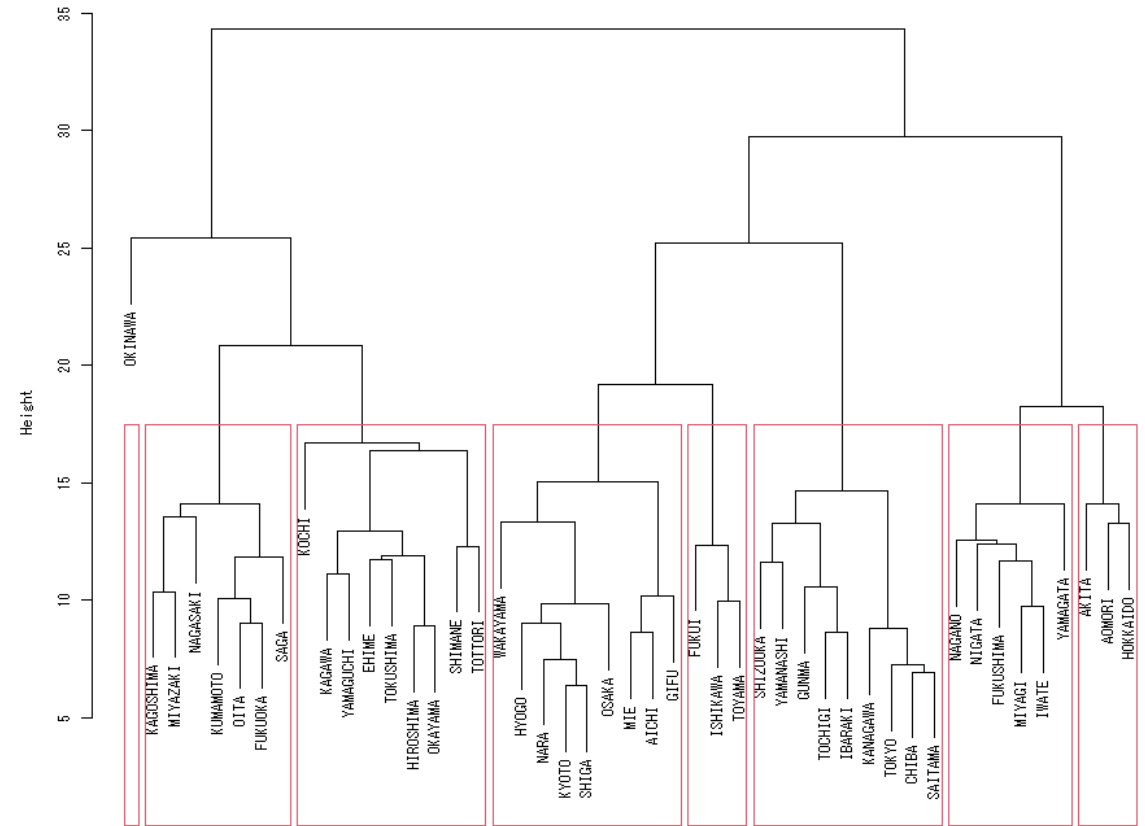


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

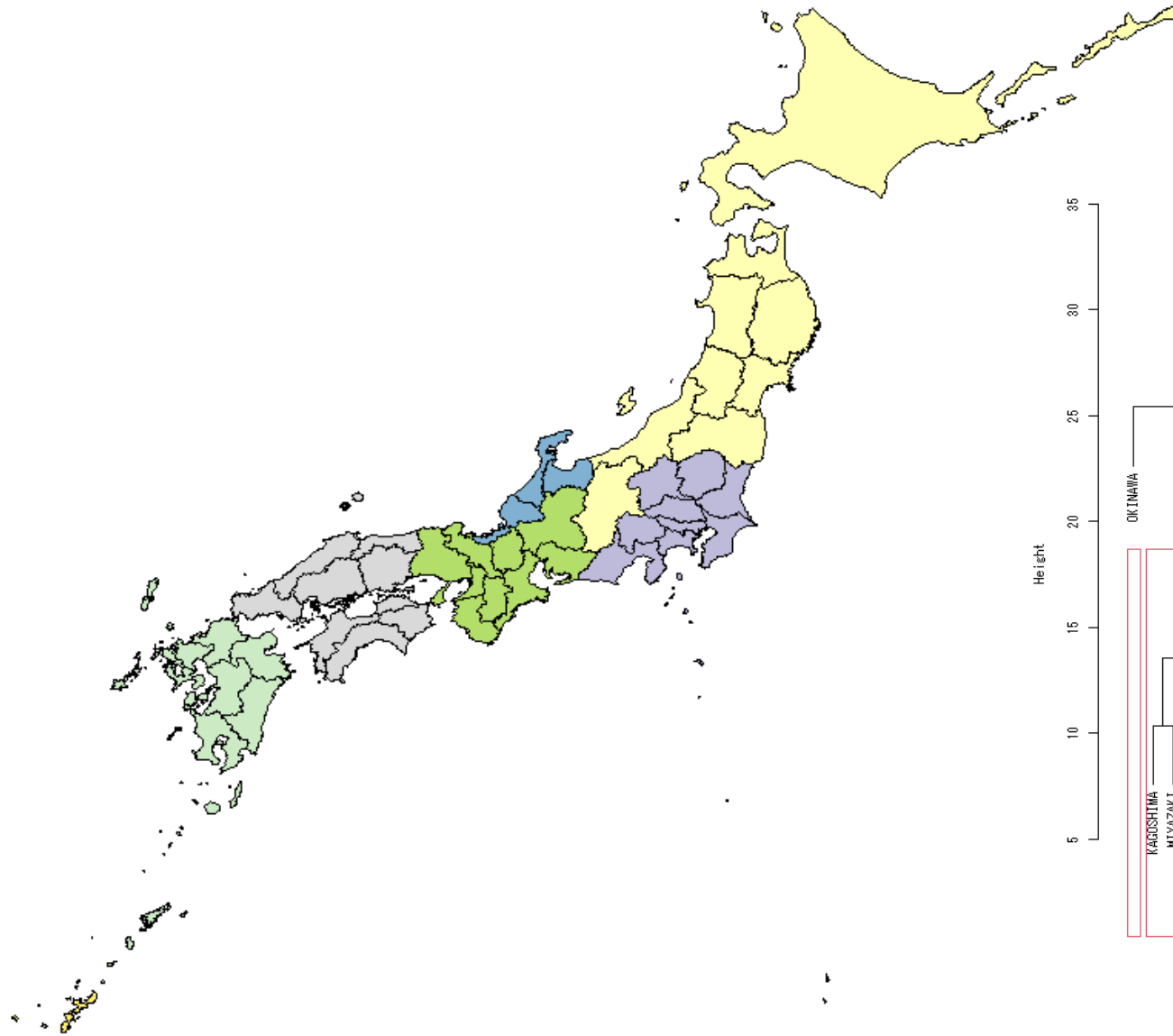


Cluster Dendrogram

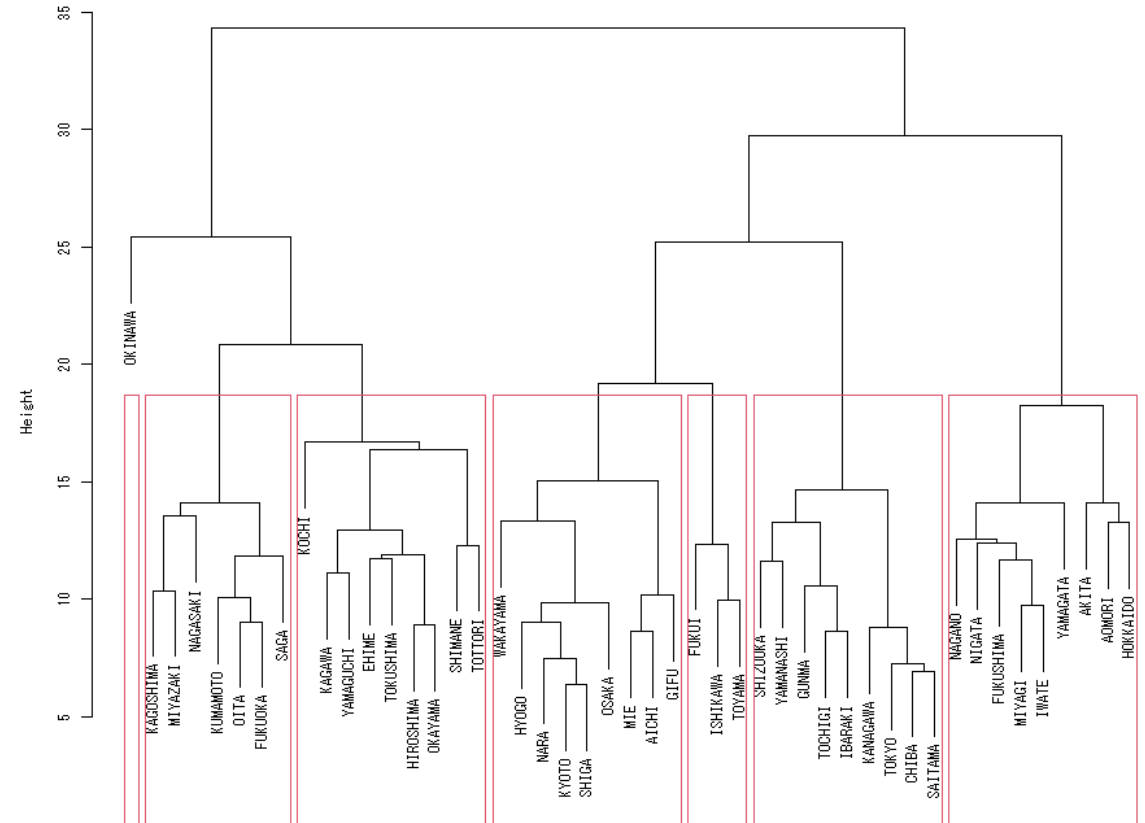


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

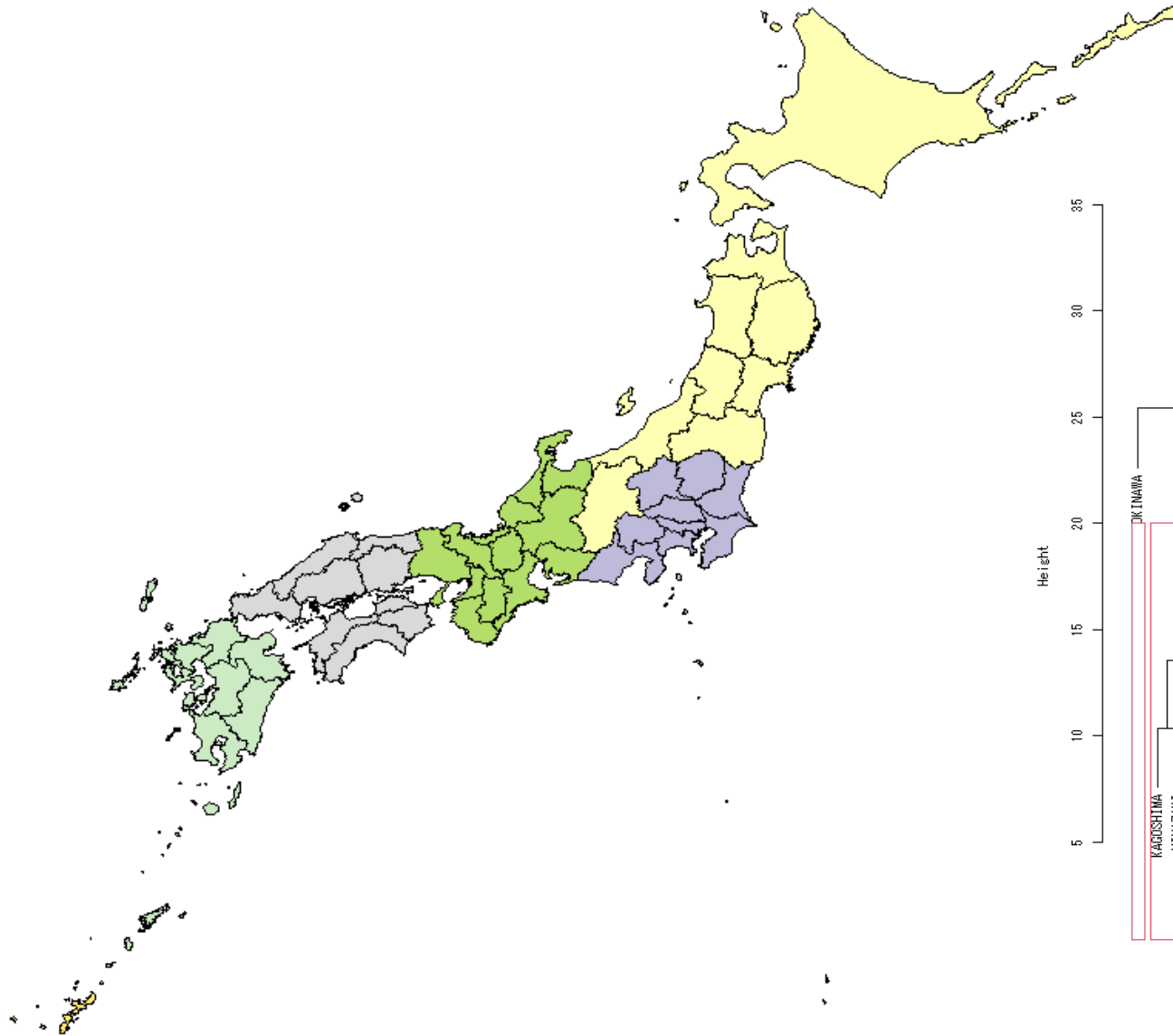


Cluster Dendrogram

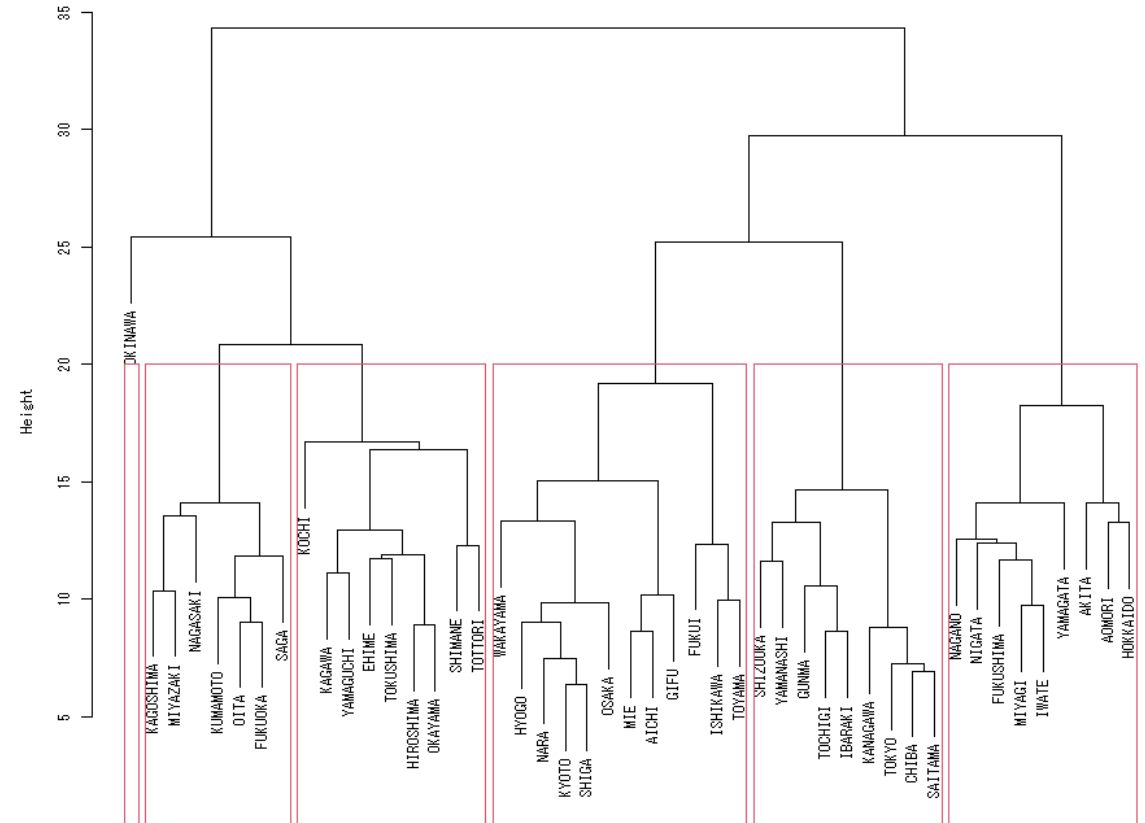


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

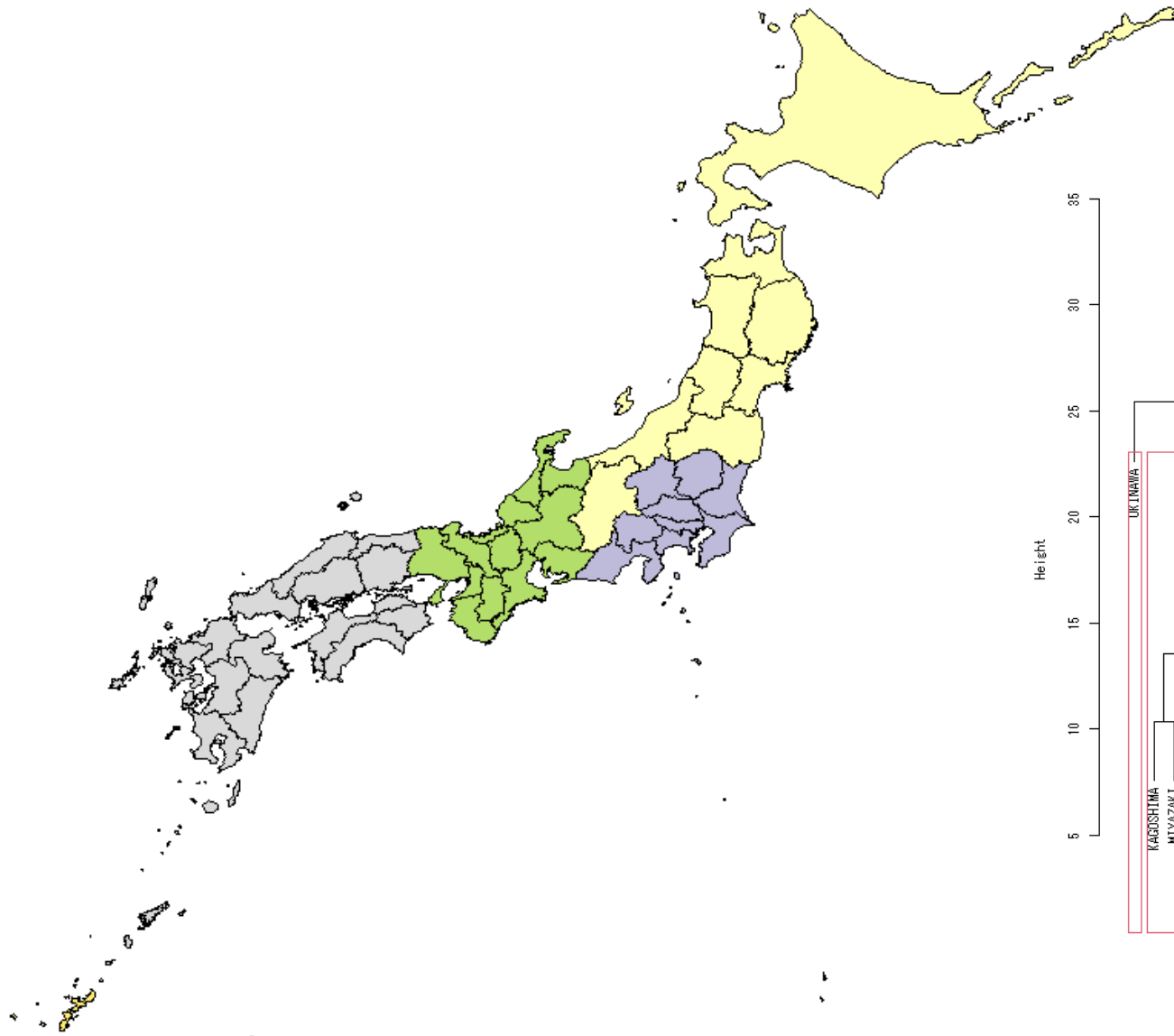


Cluster Dendrogram

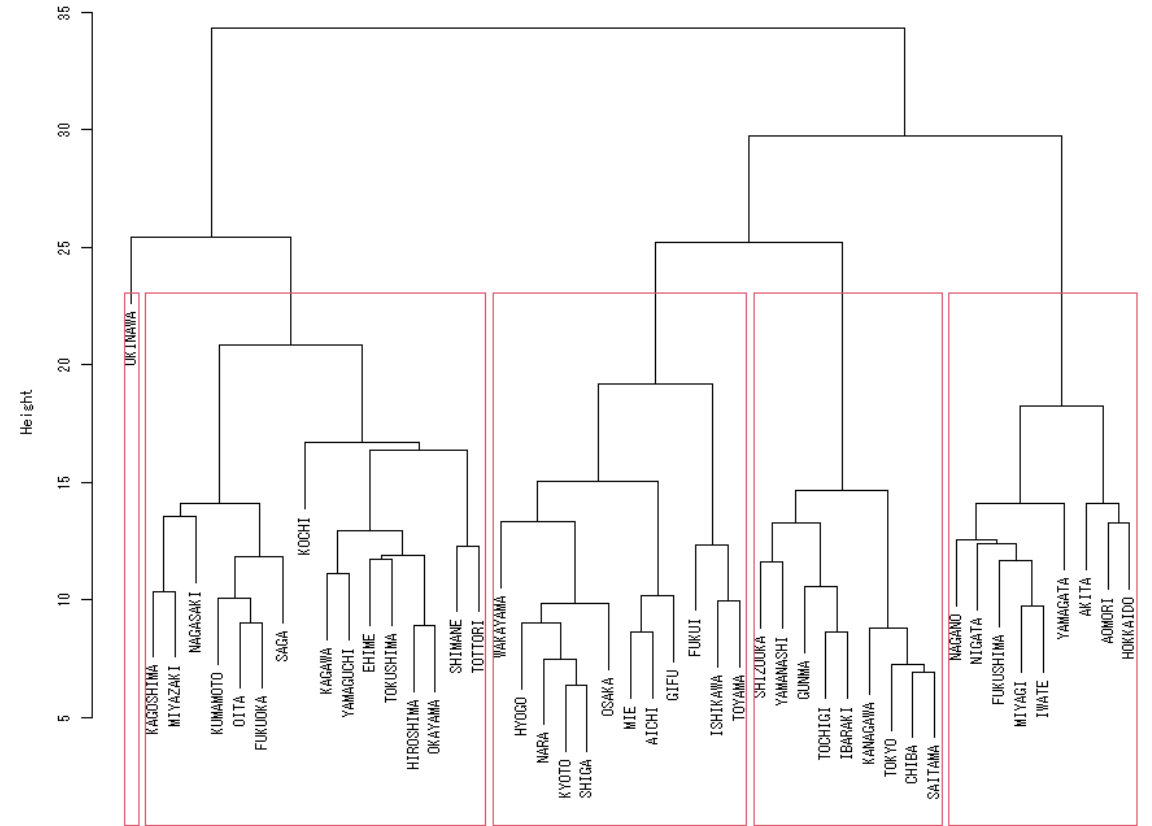


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

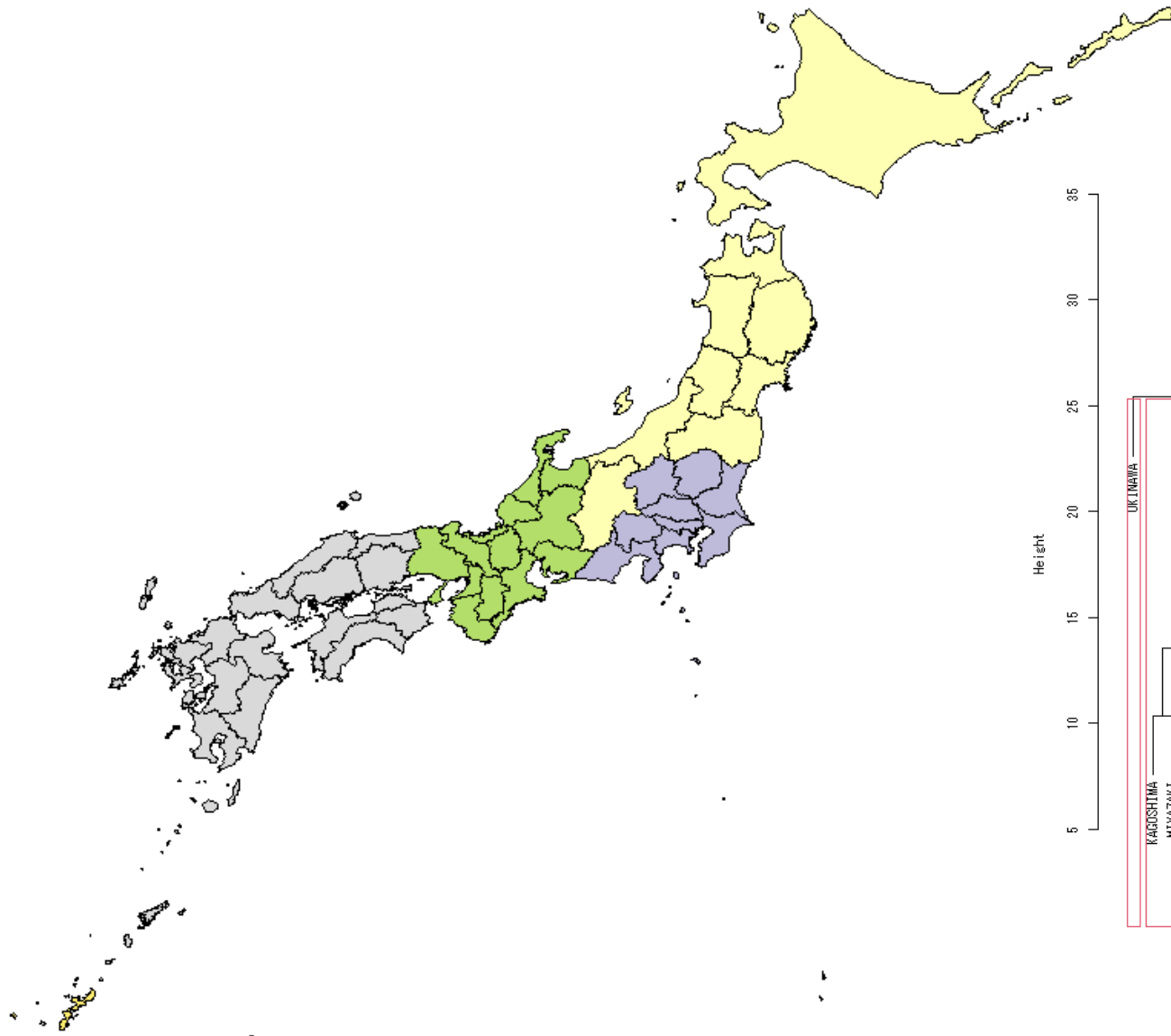


Cluster Dendrogram

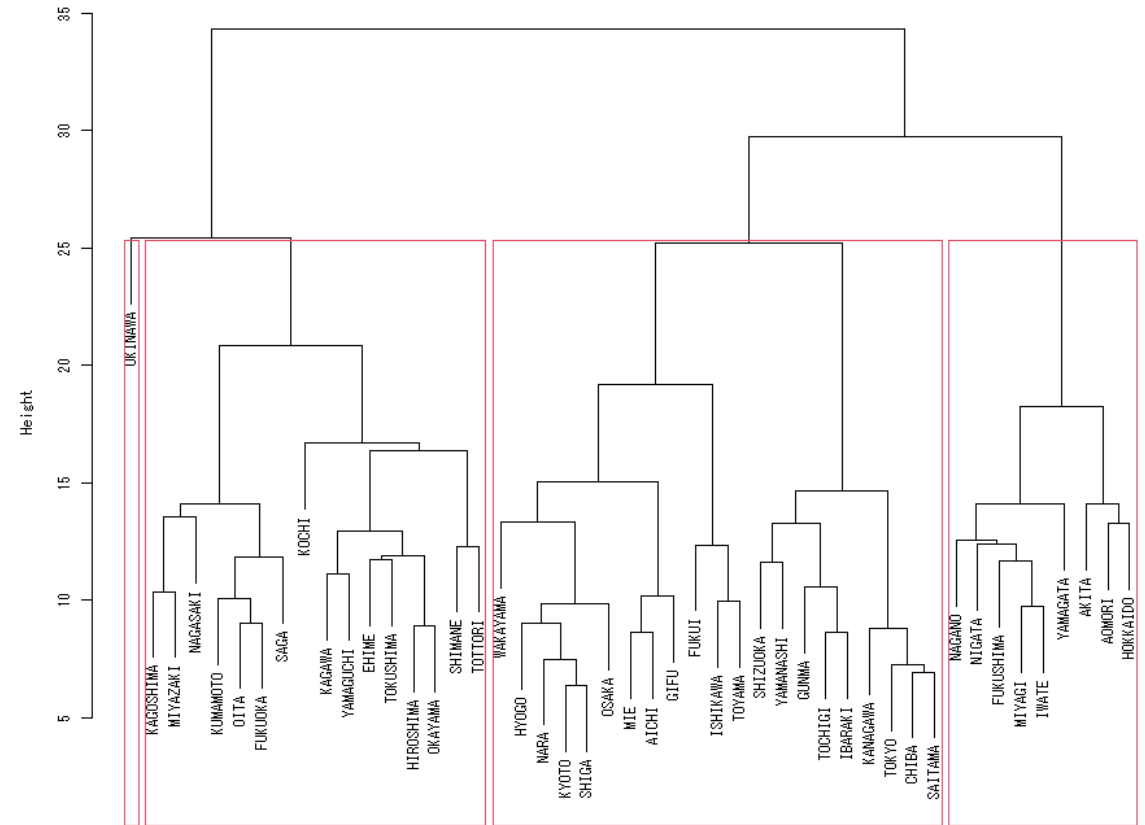


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

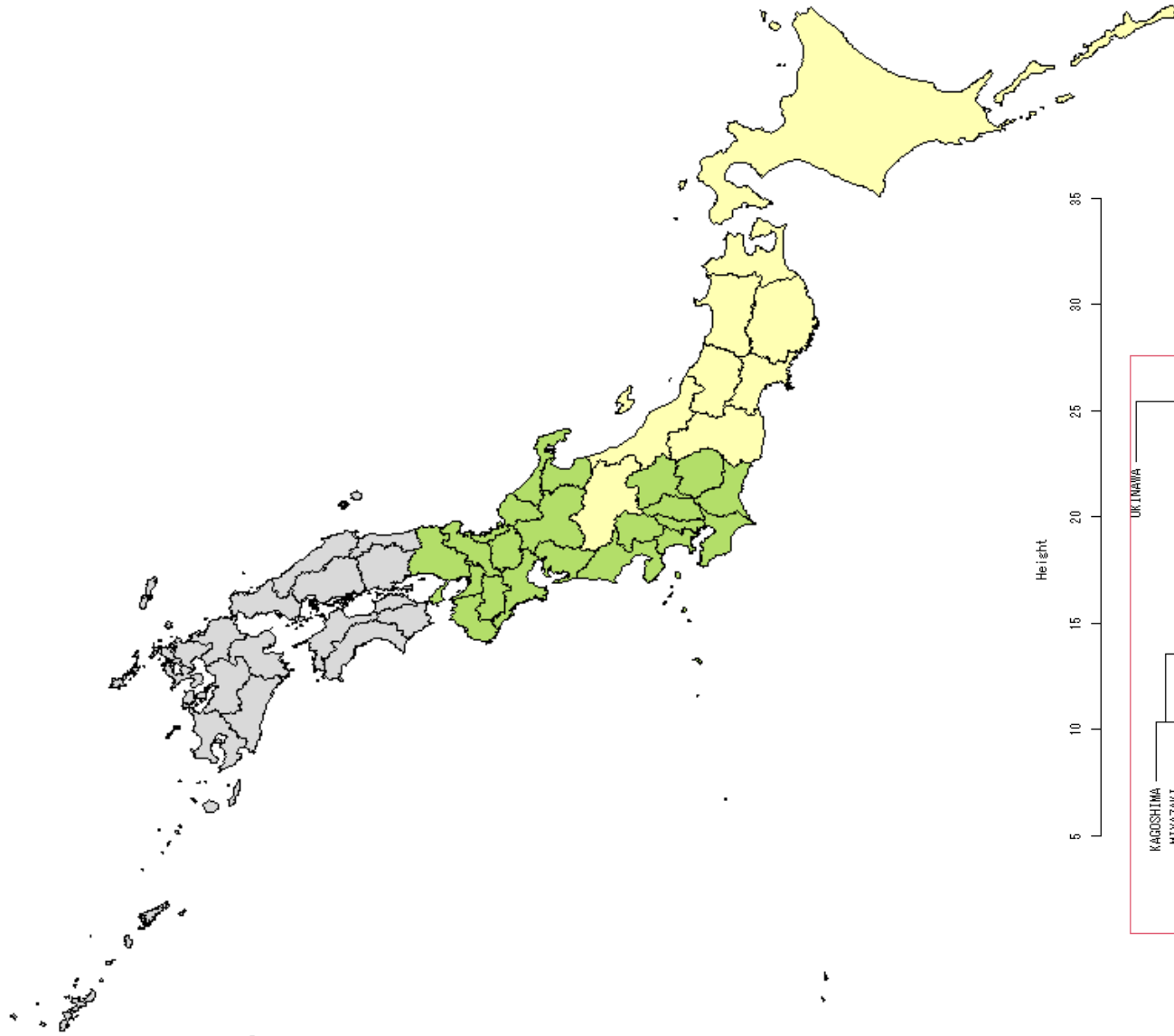


Cluster Dendrogram

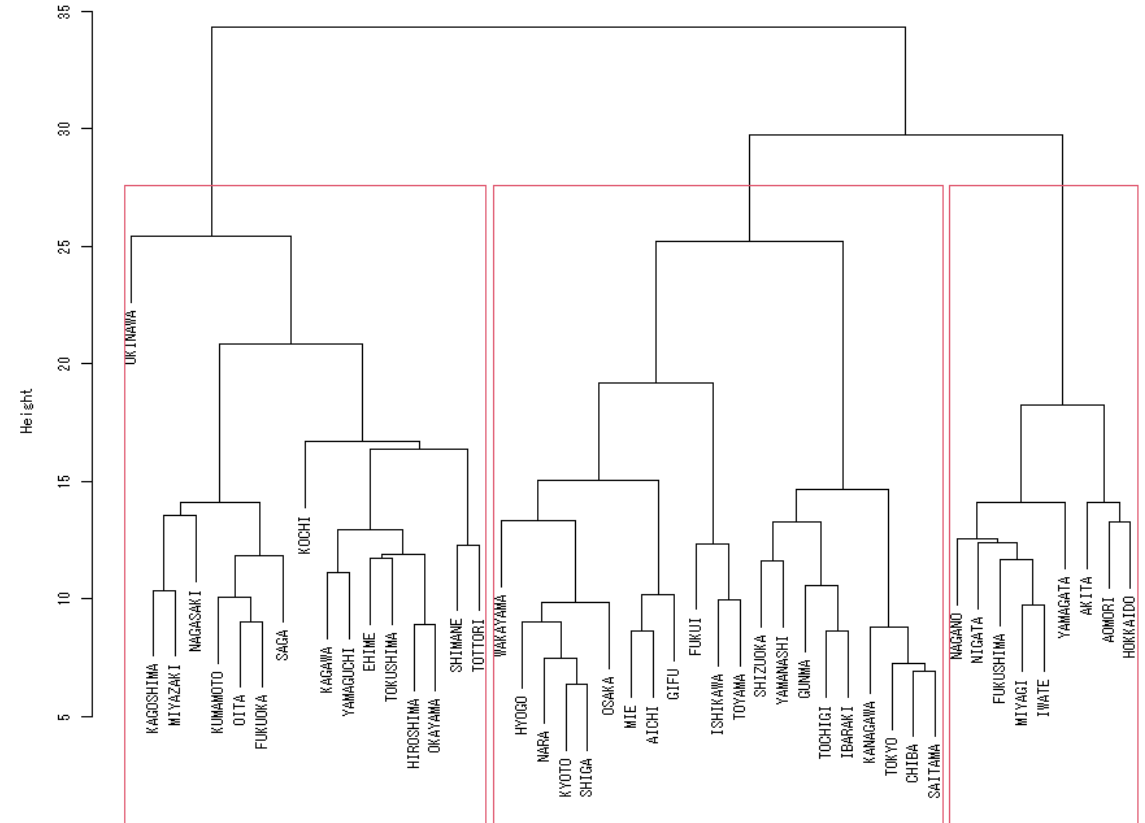


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

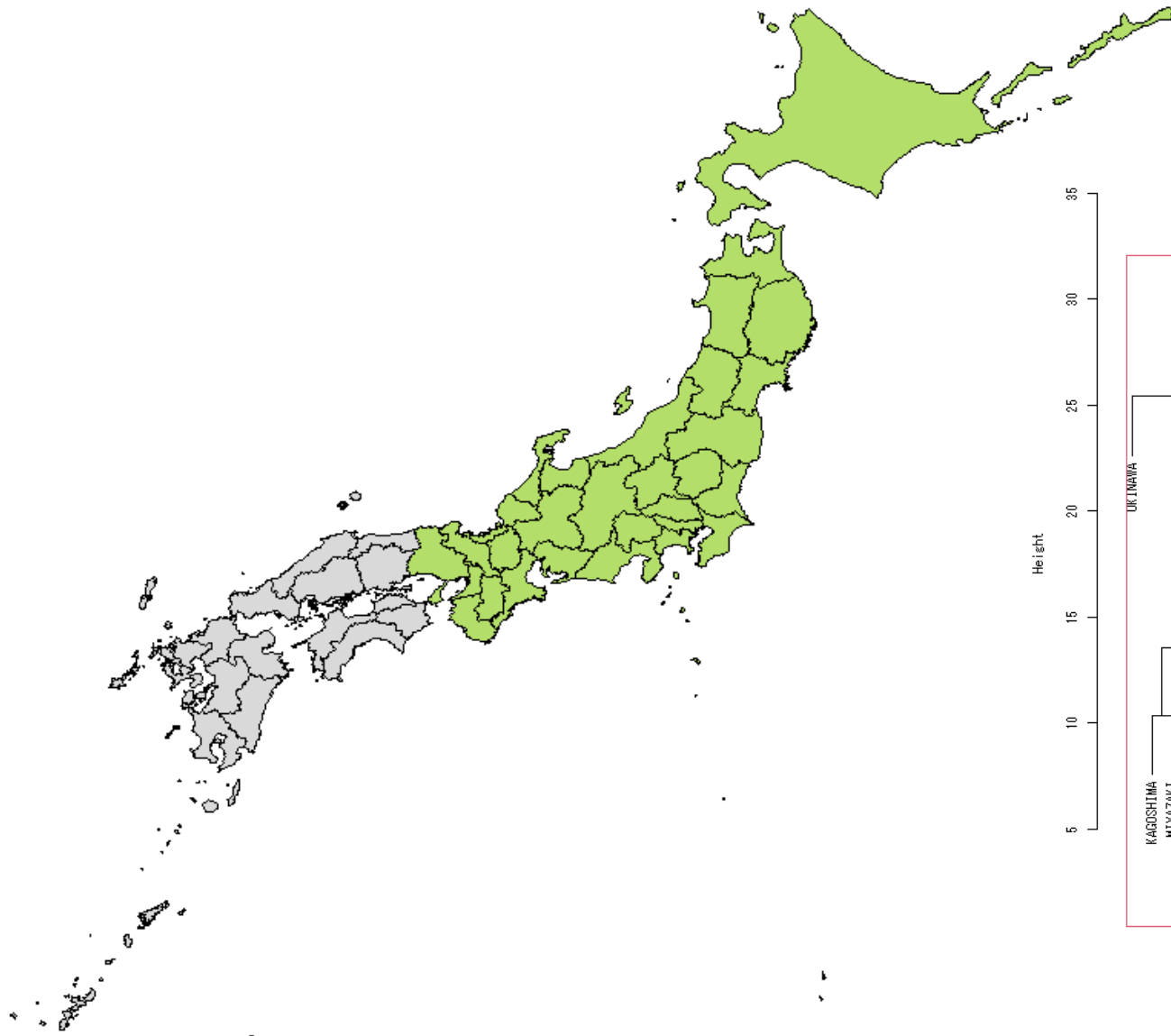


Cluster Dendrogram

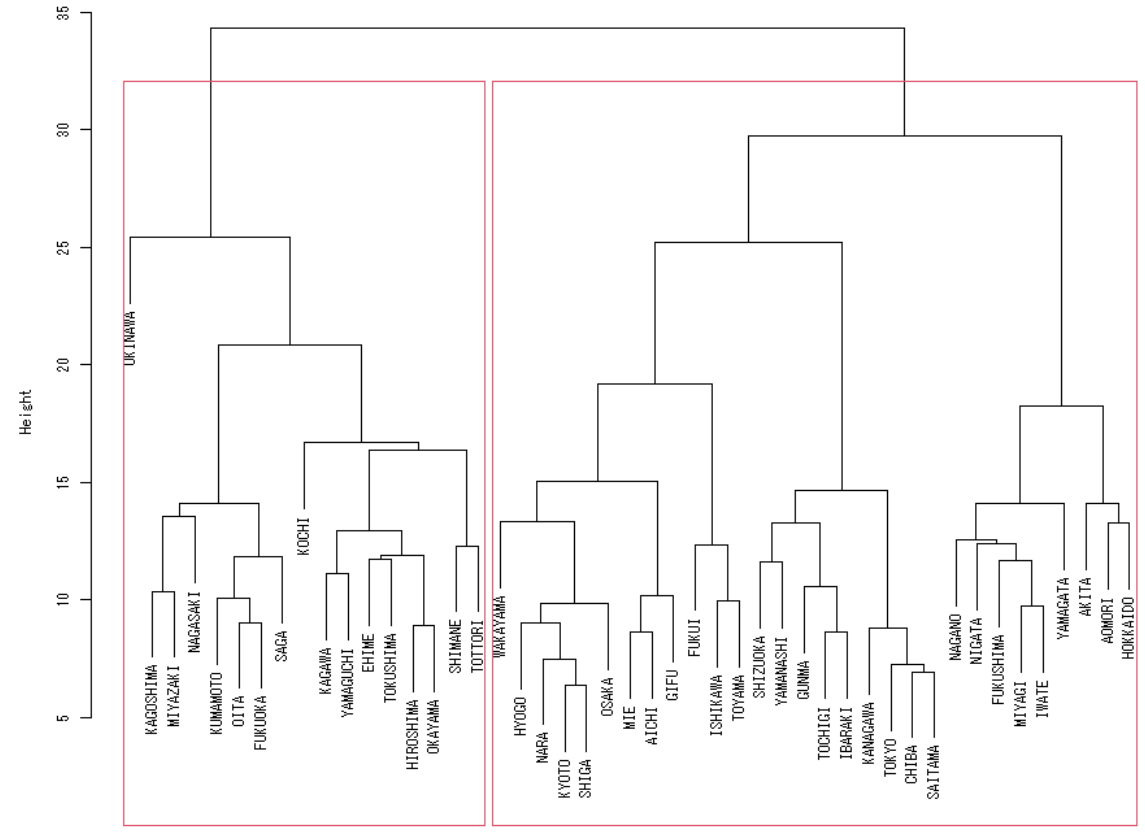


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001

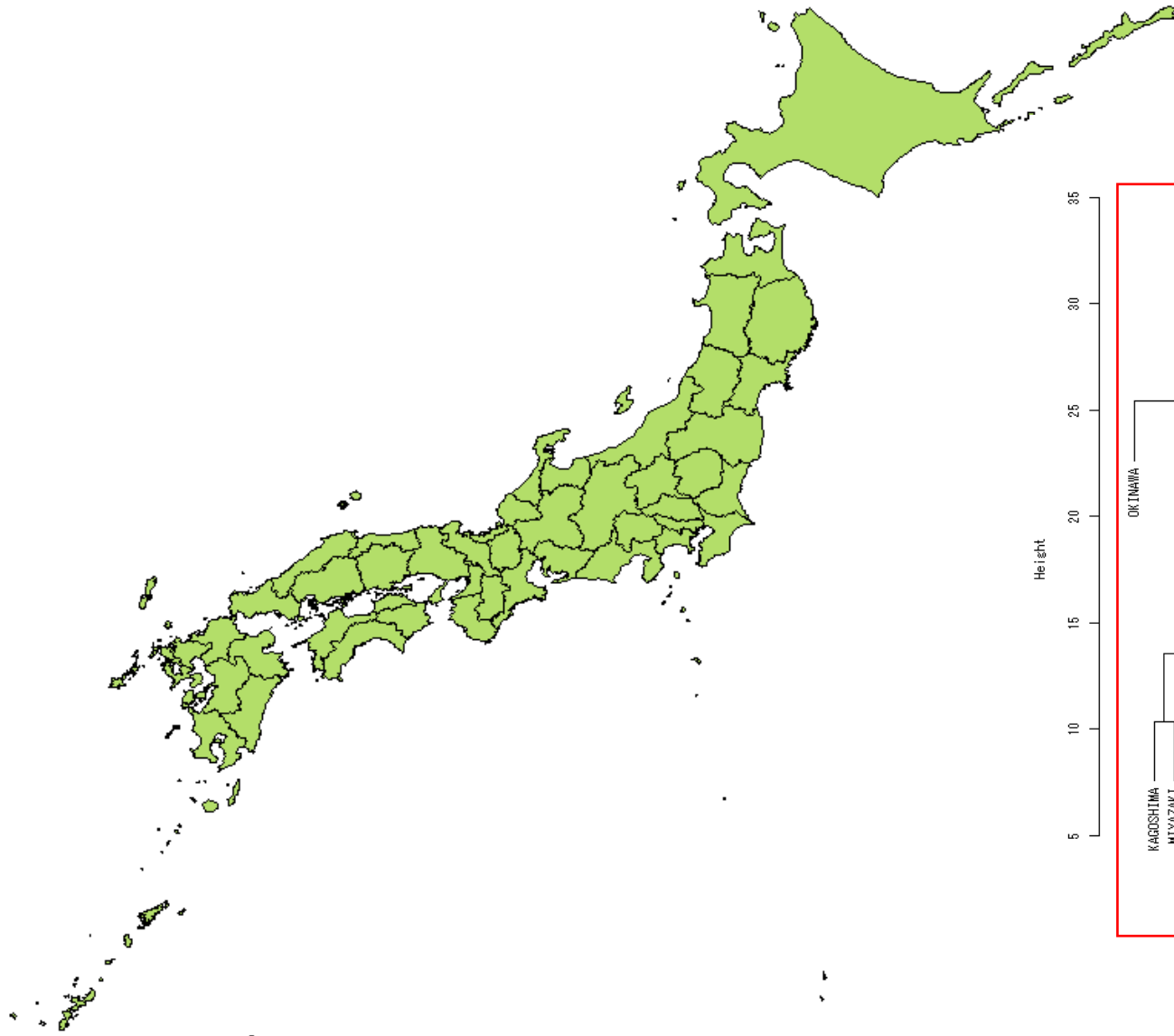


Cluster Dendrogram

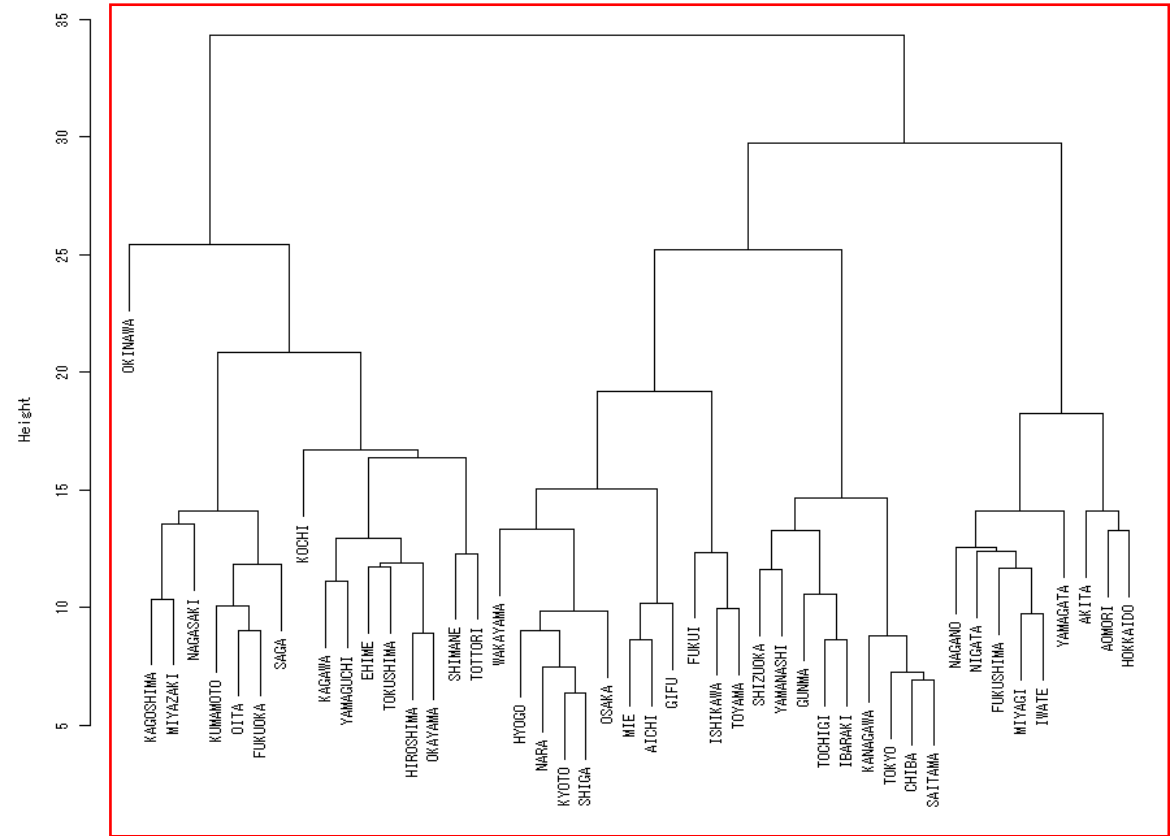


ssdse\_small.dist  
hclust (\*, "ward.D2")

N03\_001



Cluster Dendrogram



ssdse\_small.dist  
hclust (\*, "ward.D2")

# Stability of Analysis Results (3 years data)

2016~2018

N03\_001

2017~2019

N03\_001

2018~2020

N03\_001

2019~2021

N03\_001

2020~2022

N03\_001

2007~2009

N03\_001

2010~2012

N03\_001

2011~2013

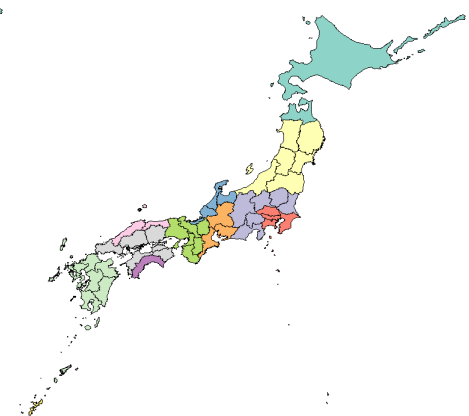
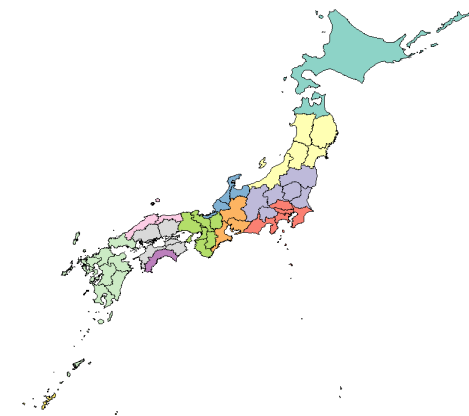
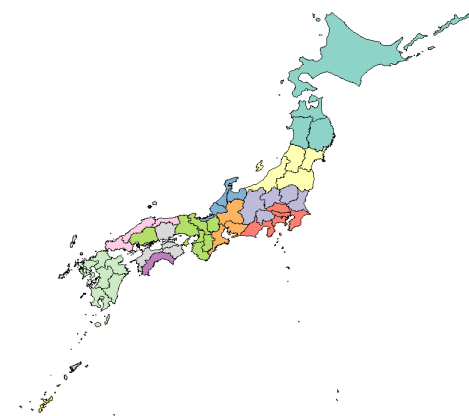
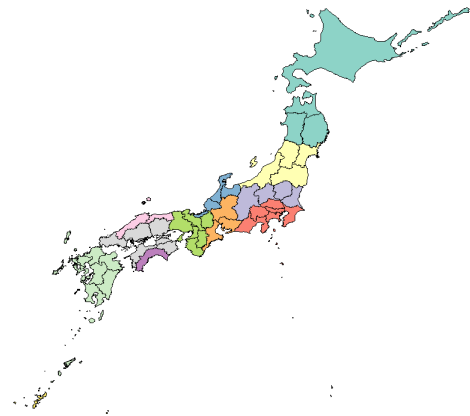
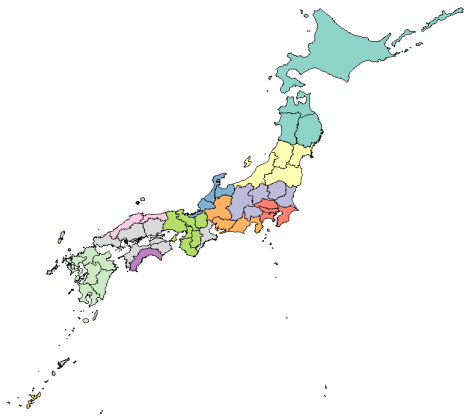
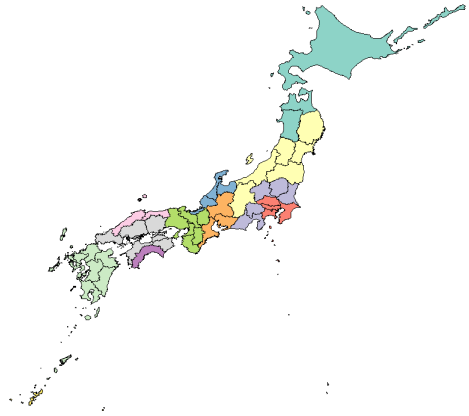
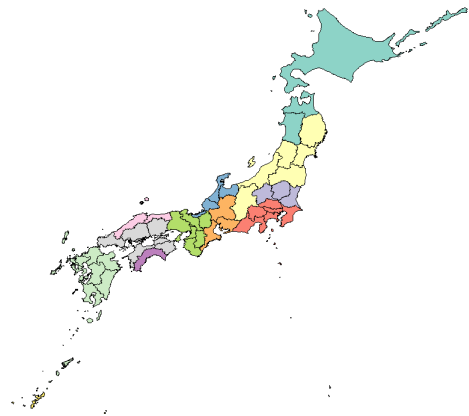
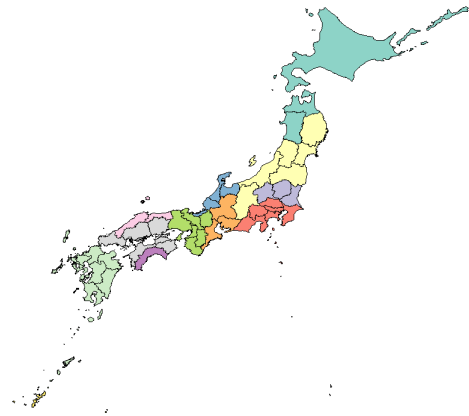
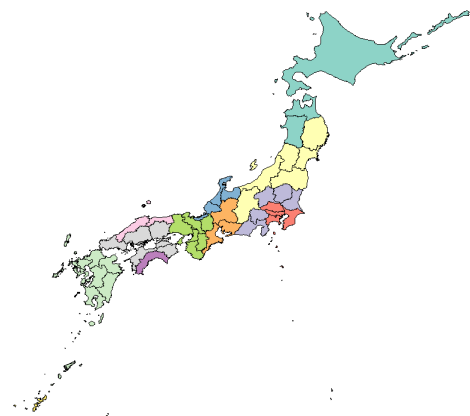
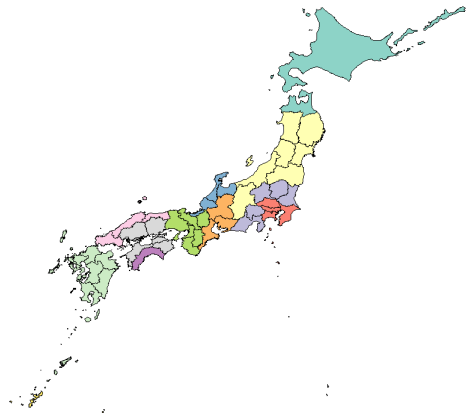
N03\_001

2012~2014

N03\_001

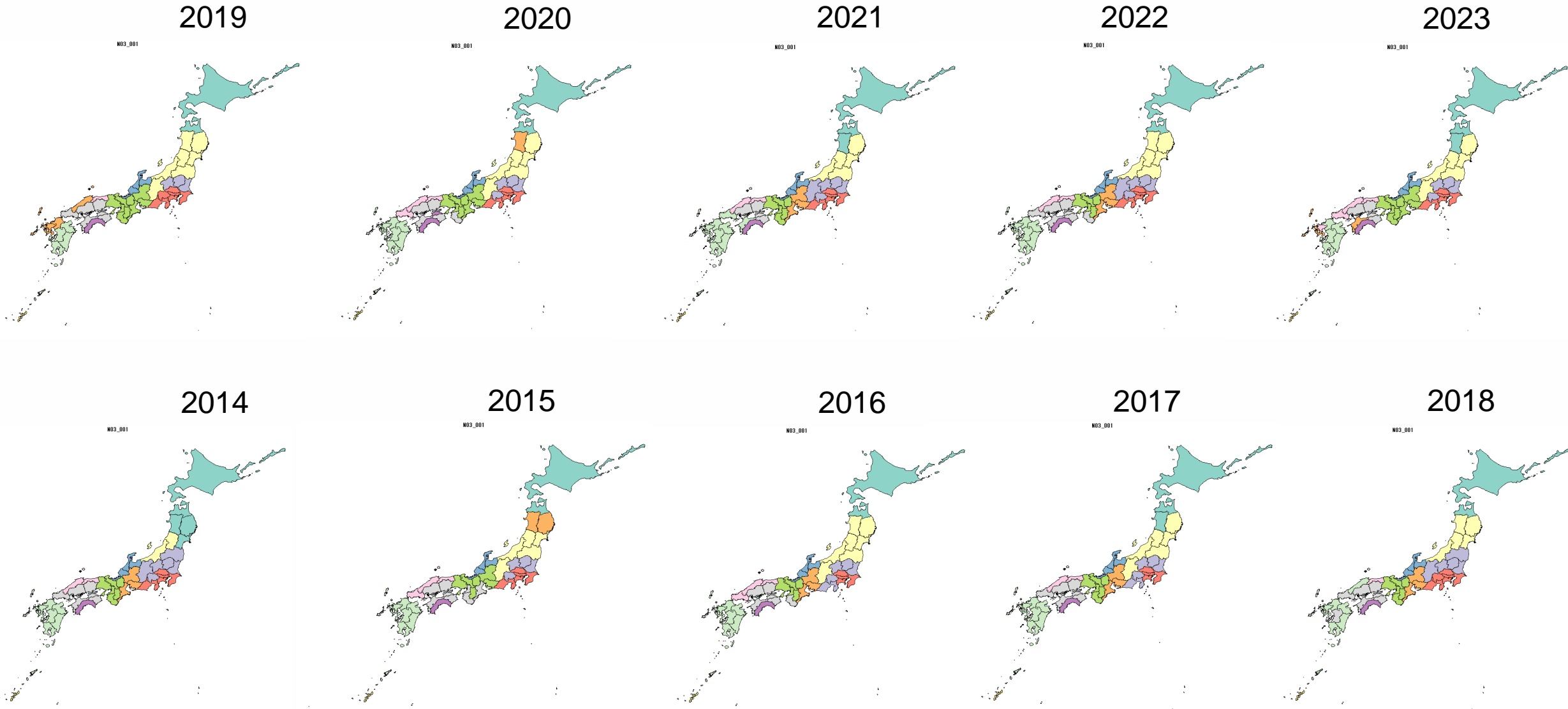
2015~2017

N03\_001



All clusters were indicated with a cluster count of 12. The results are almost the same.

# Stability of Analysis Results (1 year data)



All clusters were indicated with a cluster count of 12. The results are almost the same.

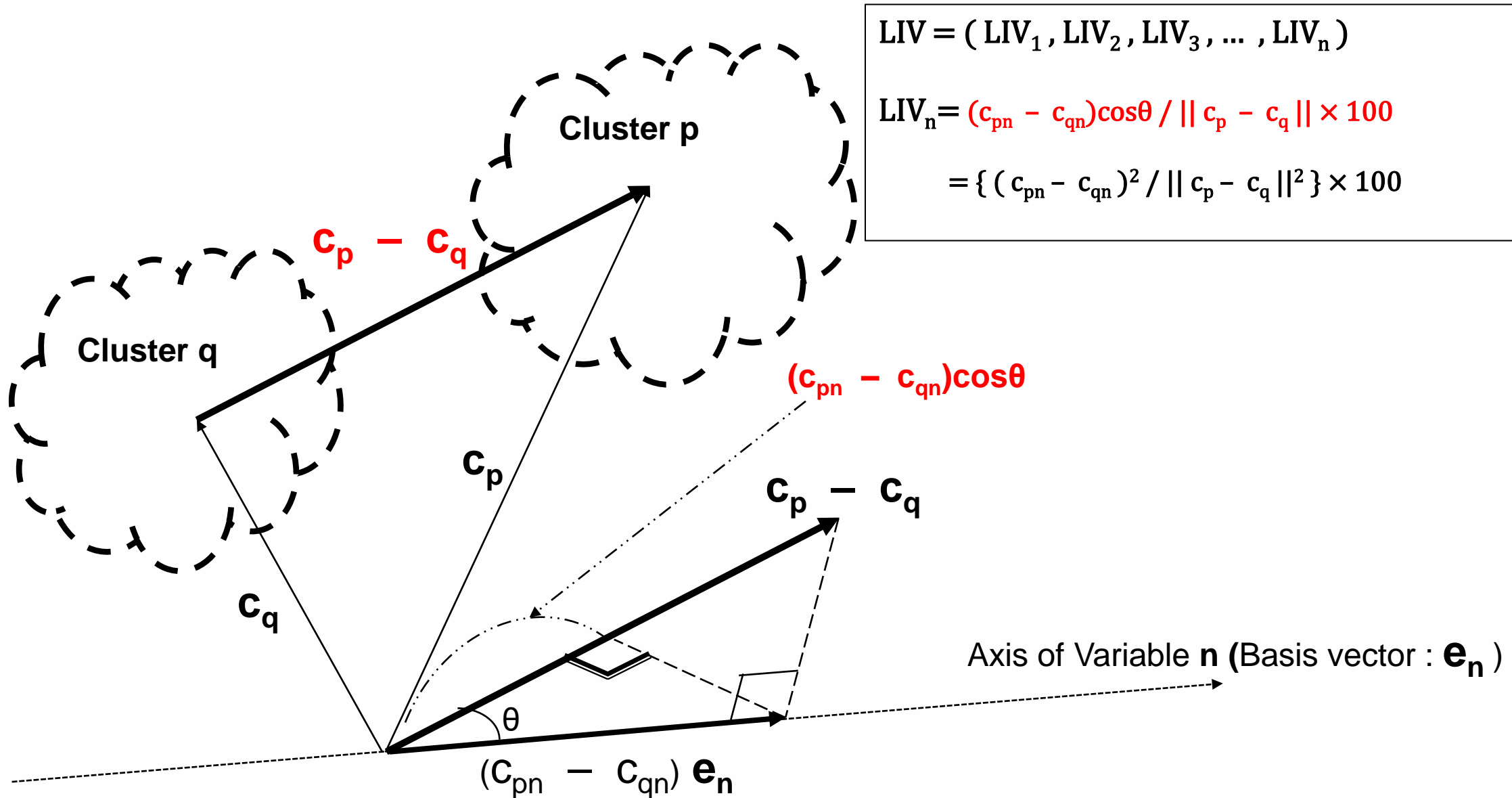
1. Analysis of Japan's "Family Income and Expenditure Survey" reveals regional characteristics of food purchasing behavior
- 2. The possibility of "GSBPM 6.2 Validate outputs" using these regional characteristics**
3. Two useful tools for comparing and evaluating the results of multiple geographic cluster analyses

## IV. Levels 1 and 2 of the GSBPM

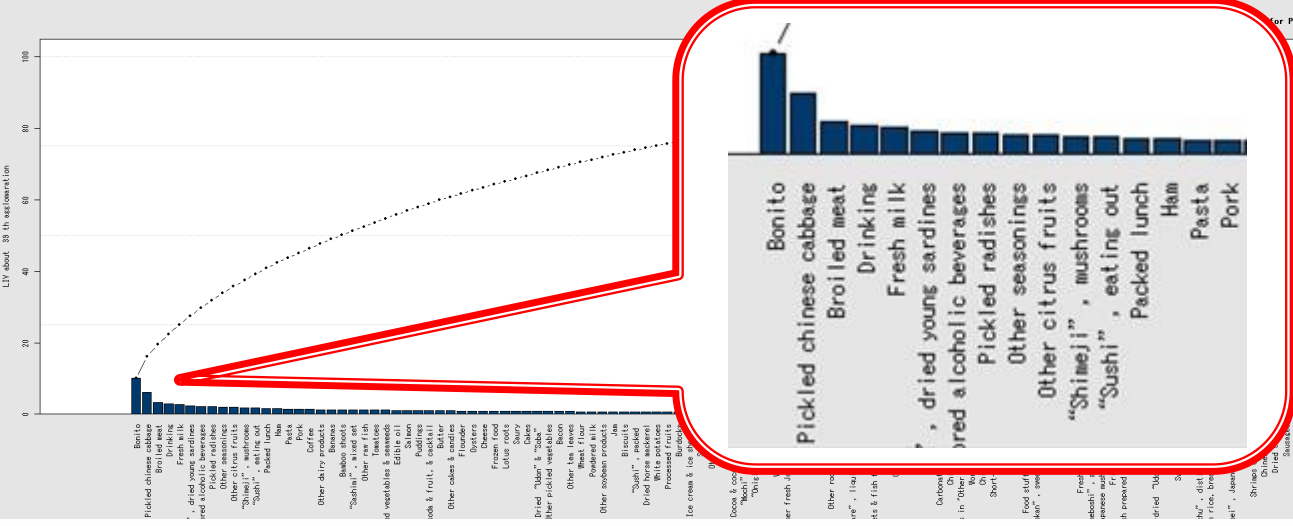
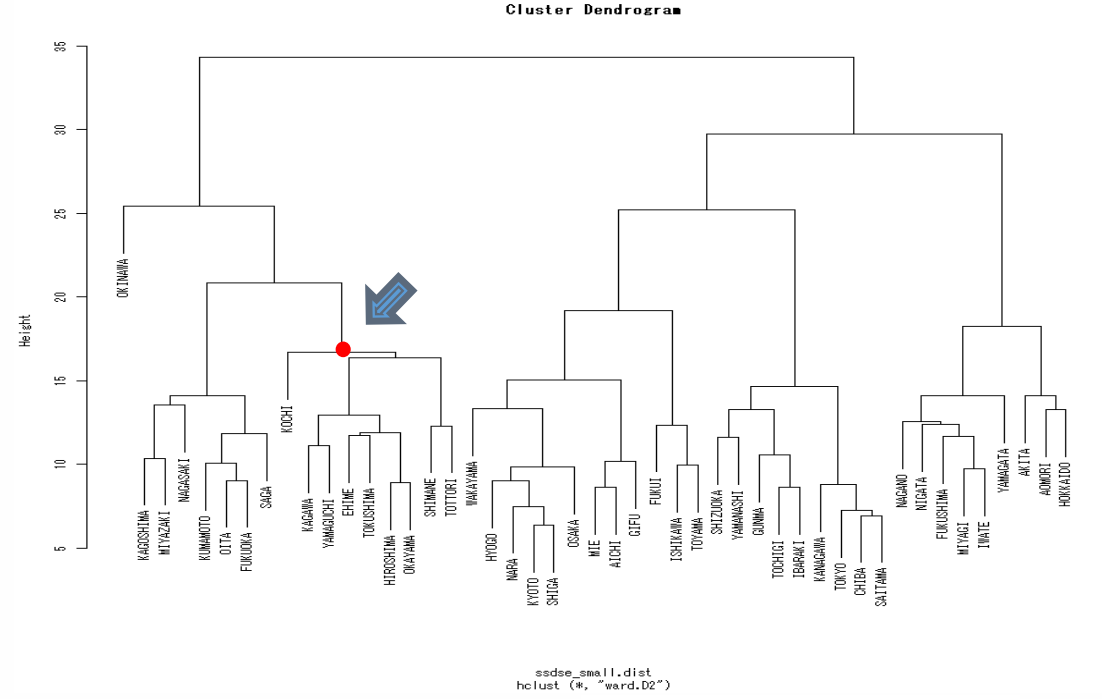
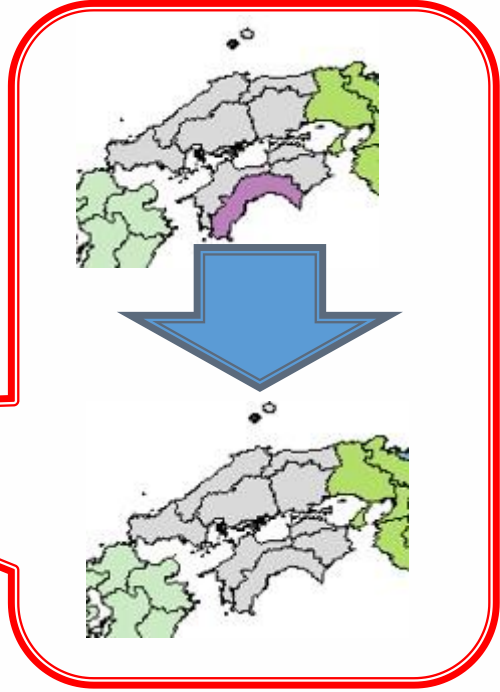
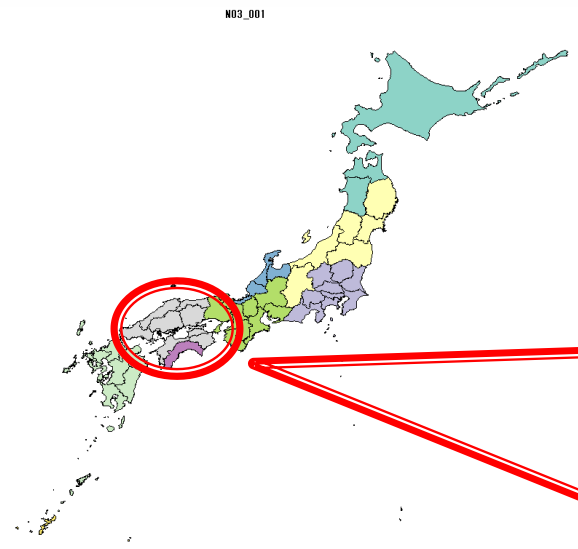
Overarching Processes							
Specify needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate
1.1 Identify needs	2.1 Design outputs	3.1 Reuse or build collection instruments	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Reuse or build processing and analysis components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Reuse or build dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalise collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production systems		5.5 Derive new variables and units	6.5 Finalise outputs	7.5 Manage user support	
1.6 Prepare and submit business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalise production systems		5.7 Calculate aggregates			
				5.8 Finalise data files			

Figure 3. The phases (level 1) and sub-processes (level 2) of the GSBPM

# Concept of “Loss of Information rate by Variable(LIV)”



# LIV pareto chart (ex. From 9 clusters to 8 clusters)



Seared bonito (Katsuo no Tataki)



1. Analysis of Japan's "Family Income and Expenditure Survey" reveals regional characteristics of food purchasing behavior
2. The possibility of "GSBPM 6.2 Validate outputs" using these regional characteristics
3. **Two useful tools for comparing and evaluating the results of multiple geographic cluster analyses**

## Providing two original functions

- i. `hclust_LIV()`
- ii. `Jaccard_identification()`

GitHub: [https://github.com/ibuchichi/R\\_function\\_2024.git](https://github.com/ibuchichi/R_function_2024.git)  
You can download R scripts I introduced today from GitHub

# LIV Calculation : hclust\_LIV()

## hclust\_LIV() function

usage:

`hclust_LIV(std_data, LIV_matrix)`

input: `std_data`

output: `LIV_matrix`

example:

`# Draw LIV pareto chart`

`Library(qcc)`

`LIV_matrix <- hclust_LIV(std_data)`

`for(i in 1: nrow(LIV_matrix)){`

`Pareto_vector <- LIV_matrix[i,]`

`names(Pareto_vector) <- colnames(LIV_matrix)`

`png_Pareto_file_name <- paste("LIV_",i ,"_pareto_qcc.png", sep="")`

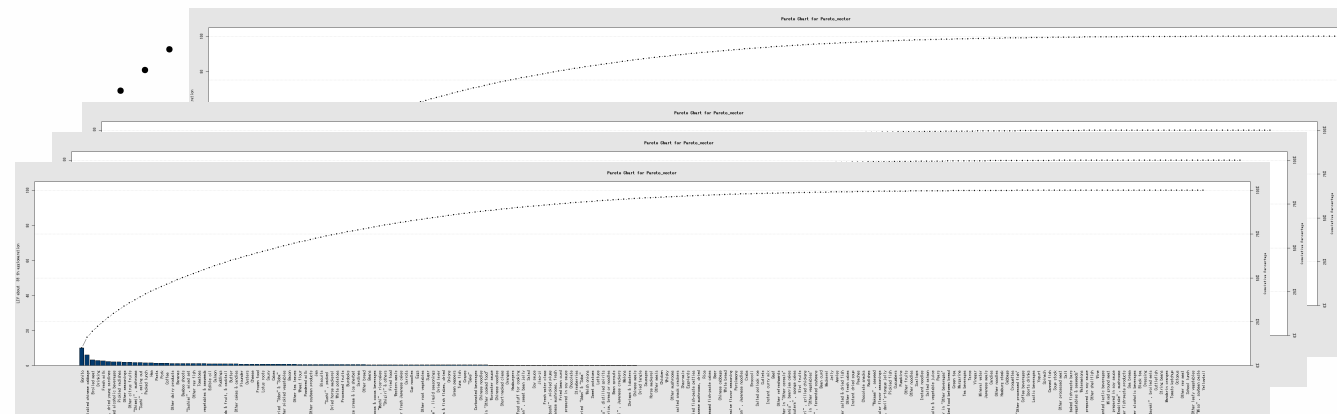
`png(png_Pareto_file_name, width=3800, height=770)`

`ylab_words <- paste("LIV about ",i , "th agglomeration")`

`pareto.chart(Pareto_vector, ylab = ylab_words)`

`dev.off()`

`}`



# Color-Coded Japan Maps

2016~2018

N03\_001

2017~2019

N03\_001

2018~2020

N03\_001

2019~2021

N03\_001

2020~2022

N03\_001

2007~2009

N03\_001

**recolored**

2010~2012

N03\_001

2011~2013

N03\_001

2012~2014

N03\_001

**recolored**

2015~2017

N03\_001

All clusters were indicated with a cluster count of 12. The results are almost the same.

# Original Color-Coded Japan Maps

2016~2018

N03\_001

2017~2019

N03\_001

2018~2020

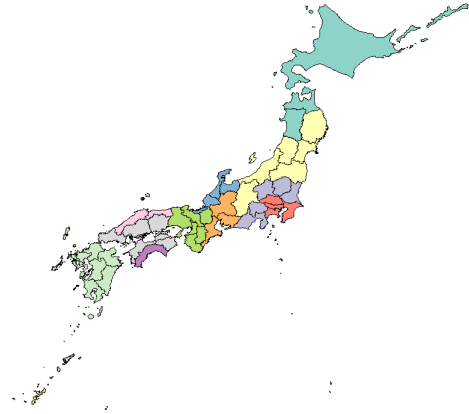
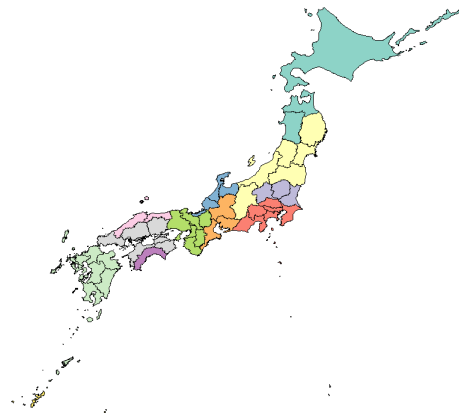
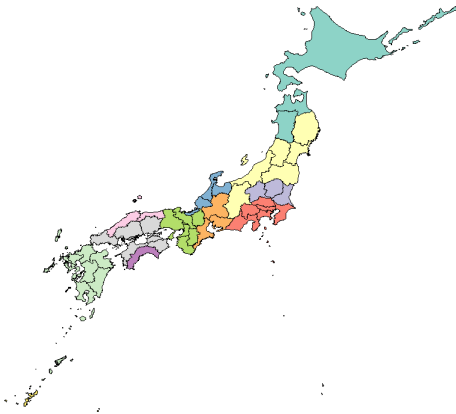
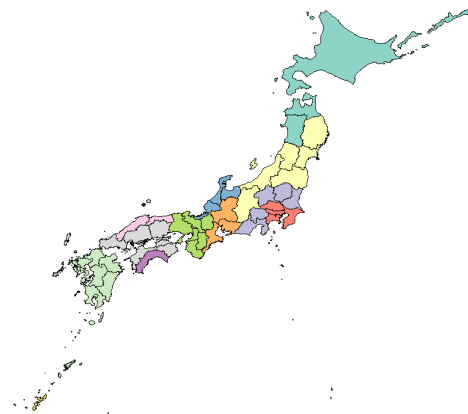
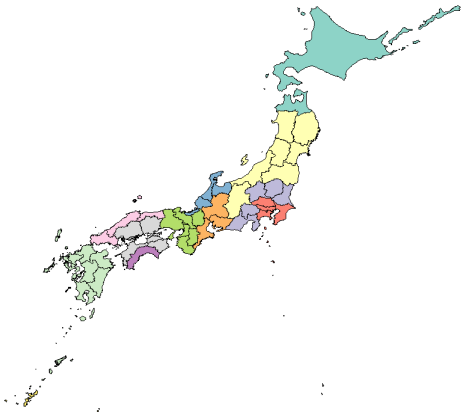
N03\_001

2019~2021

N03\_001

2020~2022

N03\_001



2007~2009

2010~2012

N03\_001

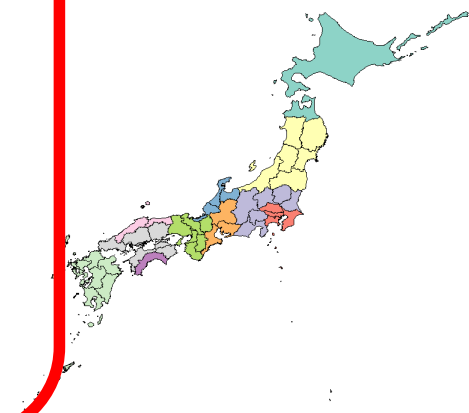
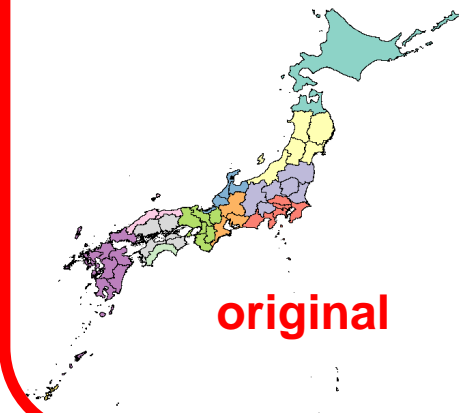
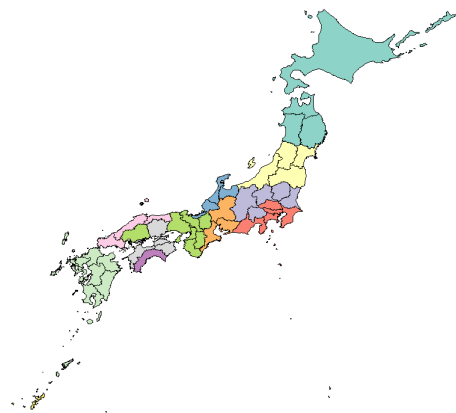
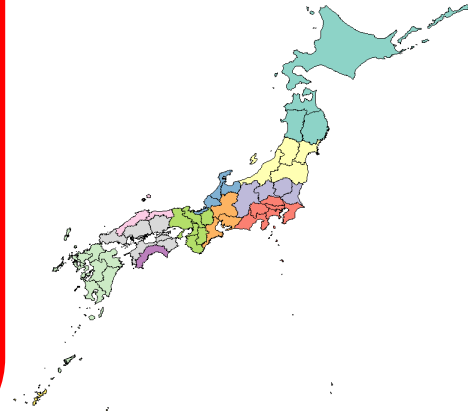
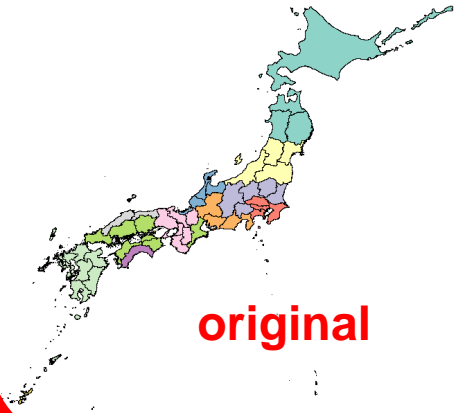
2011~2013

N03\_001

2012~2014

2015~2017

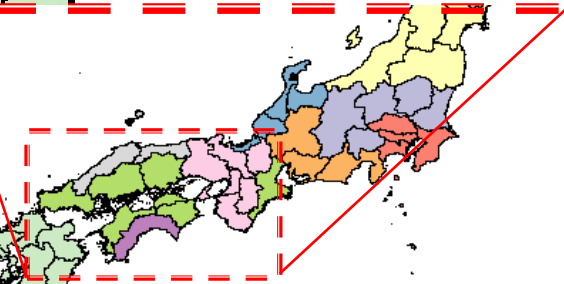
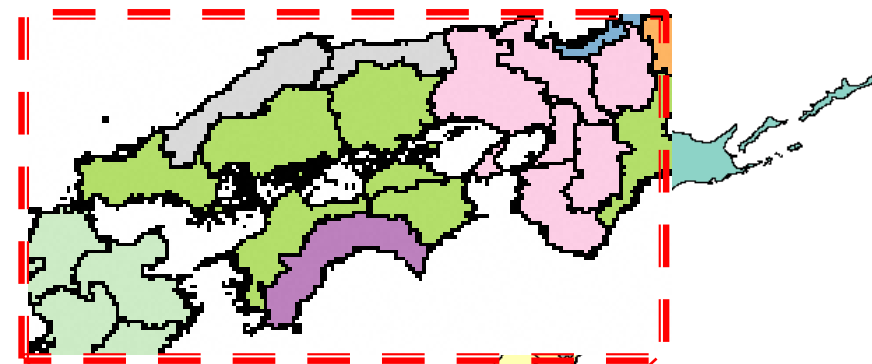
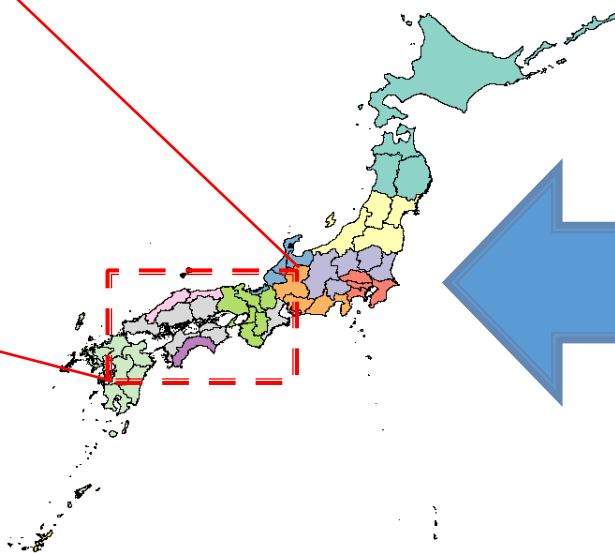
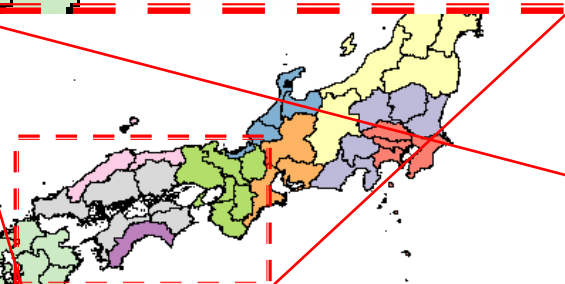
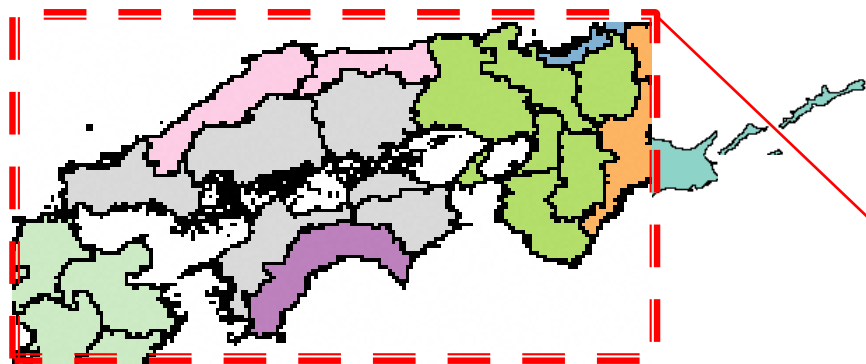
N03\_001



All clusters were indicated with a cluster count of 12. The results are almost the same.

# Jaccard\_identification() Function

Similar clusters are not necessarily the same color !



**Reference Map**  
2017-2019

2007-2009  
**recolored**  
by Jaccard\_identification()  
function

2007-2009  
**original**

# Draw Recolored Map : Jaccard\_identification()

**Jaccard\_identification()** function

usage:

```
Jaccard_identification(Ref_hclust.cnum, hclust.cnum )
```

```
input: Ref_hclust.cnum, hclust.cnum
```

```
output: Recolor_hclust.cnum
```

example:

```
# Draw Japan Map
```

```
Ref_hclust.cnum <- cutree(Ref_hclust.result, k=12)
```

```
hclust.cnum <- cutree(hclust.result, k=12)
```

```
Recolor_hclust.cnum <- Jaccard_identification(Ref_hclust.cnum, hclust.cnum)
```

```
library(NipponMap)
```

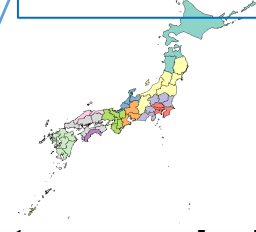
```
library(RColorBrewer)
```

```
NipponMap::JapanPrefMap(brewer.pal(12,"Set3")[Ref_hclust.cnum])
```

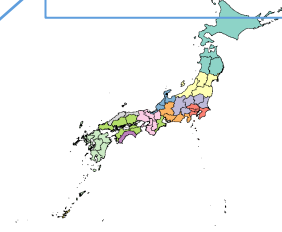
```
NipponMap::JapanPrefMap(brewer.pal(12,"Set3")[hclust.cnum])
```

```
NipponMap::JapanPrefMap(brewer.pal(12,"Set3")[Recolor_hclust.cnum])
```

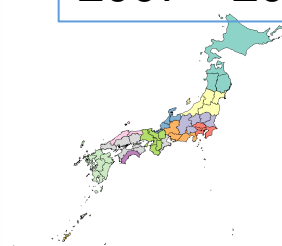
2017~2019 Reference



2007~2009 original



2007~2009 recolored



- The set  $Y$  consisting of  $N$  sources is divided into  $K$  clusters.

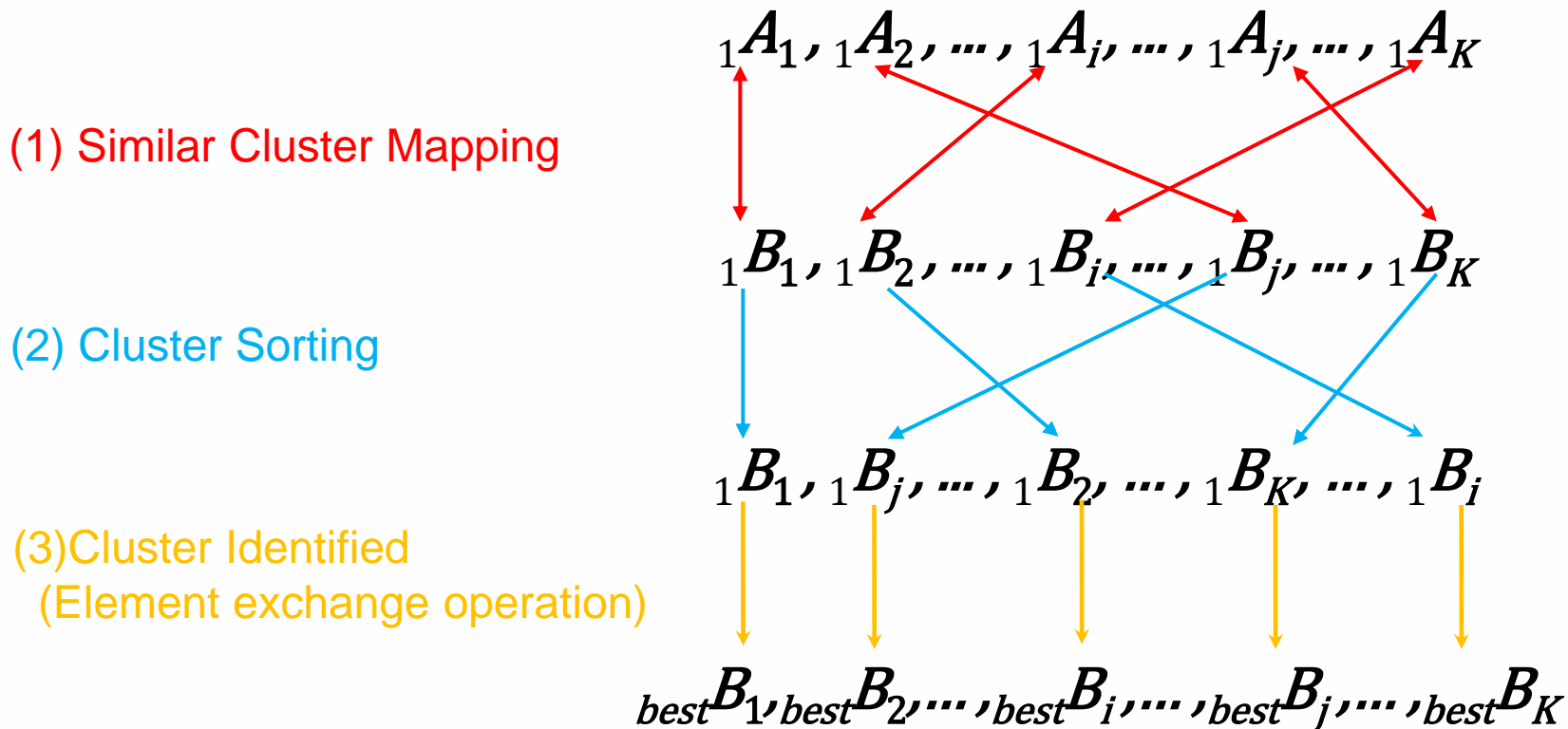
$$Y = \{ 1, 2, 3, \dots, i, \dots, j, \dots, N \}$$

Each cluster is generated by survey data a and b with differing survey years.

Let  ${}_1A = \{ {}_1A_1, {}_1A_2, \dots, {}_1A_i, \dots, {}_1A_j, \dots, {}_1A_K \}$  be the partitions based on data a

Let  ${}_1B = \{ {}_1B_1, {}_1B_2, \dots, {}_1B_i, \dots, {}_1B_j, \dots, {}_1B_K \}$  be the partitions based on data b

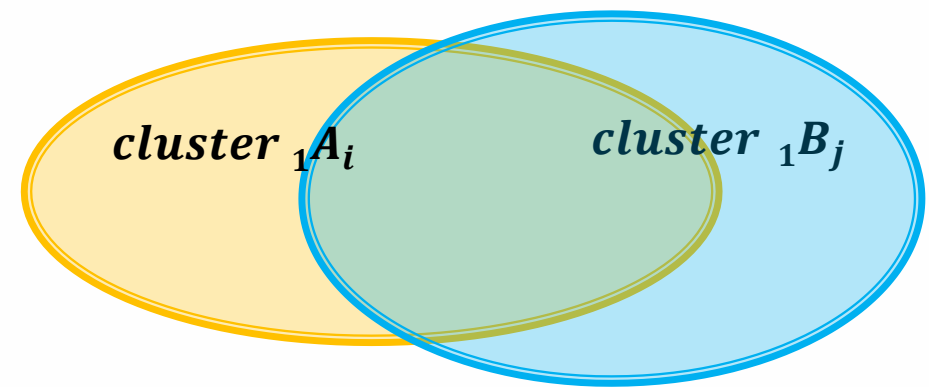
$$Y = \bigcup_{i=1}^K {}_1A_i = \bigcup_{i=1}^K {}_1B_i = \bigcup_{i=1}^K {}_m B_i \quad \text{where } {}_m B (1 \leq m \leq K!): {}_1B \text{ permutation}$$



Jaccard Cluster Similarity Matrix(  $JAC$  ) is defined below:

$$JAC_{ij} = \frac{[{}_1A_i \cap {}_1B_j]}{[{}_1A_i \cup {}_1B_j]}$$

$$JAC = \begin{pmatrix} JAC_{11} & \cdots & JAC_{1K} \\ \vdots & \ddots & \vdots \\ JAC_{K1} & \cdots & JAC_{KK} \end{pmatrix}$$



**Matching Algorithm:**

Assume that

$$JAC_{pq} = \max_{1 \leq i \leq K, 1 \leq j \leq K} JAC_{ij}. \quad (\text{Equation 1})$$

$\Rightarrow$  **cluster  ${}_1A_p$**  and **cluster  ${}_1B_q$**  are the most similar clusters.

$\Rightarrow$  Match **cluster  ${}_1A_p$**  with **cluster  ${}_1B_q$**  .

Replace  $(JAC_{p1}, \dots, JAC_{pK})$  and  $(JAC_{1q}, \dots, JAC_{Kq})$  in  $JAC$  with  $(0, \dots, 0)$

$JAC$  itself will change each time.

(The procedure for matching for clusters other than the already matched **cluster  ${}_1A_p$**  and **cluster  ${}_1B_q$**  .)

Again determine similar clusters by Equation 1. Repeat this operation until all clusters are matched.

- 1. We have shown that cluster analysis of Japanese “Family income and expenditure survey” clearly reveals regional differences in food purchasing behavior.**
- 2. We discussed possibility can be used for “GSBPM 6.2 Validate outputs”. And then we proposed "LIV method".**
- 3. Finally, we introduced two original R functions for comparing the results of multiple cluster analyses.**

# Thank you so much!



Atsushi Kimura  
e-mail: [atkimura@nstac.go.jp](mailto:atkimura@nstac.go.jp)

GitHub: [https://github.com/ibuchichi/R\\_function\\_2024.git](https://github.com/ibuchichi/R_function_2024.git)  
You can download R scripts and Data Sets I introduced today from GitHub