



JADAA

オルタナティブデータ推進協議会

Alternative Data Factbook 2024

“Pace of growth in data usage has slowed but become more diversified”

(Digest)

February, 2025

Japan Alternative Data Accelerator Association

Sompo Institute Plus Inc.



Survey Overview

Issue to explore

- Foreign countries have been building up knowledge and insights on alternative data, and their research and studies on data use have also been progressing.
- In Japan, there is not sufficient analysis on the use of alternative data, such as "who uses the data", "what type of data is used", or "how often the data is used". To understand the present situations and issues surrounding alternative data would promote the use of alternative data.

About this survey

- To obtain a clear understanding about the use of alternative data in Japan and issues involved in it, we have conducted a survey of members of the Japan Alternative Data Accelerator Association (this is the third survey).
- We compiled this Factbook from answers provided by 56 respondents.

■ Survey respondents

| Business | Number of Respondents | Ratio |
|--|-----------------------|-------|
| Asset management | 5 | 9% |
| Finance and insurance (excluding asset management) | 7 | 13% |
| Manufacturing | 2 | 4% |
| Think tank | 3 | 5% |
| Education | 7 | 13% |
| Information technology/System development | 21 | 38% |
| Other | 11 | 20% |
| Total | 56 | 100% |

■ Survey data collection period

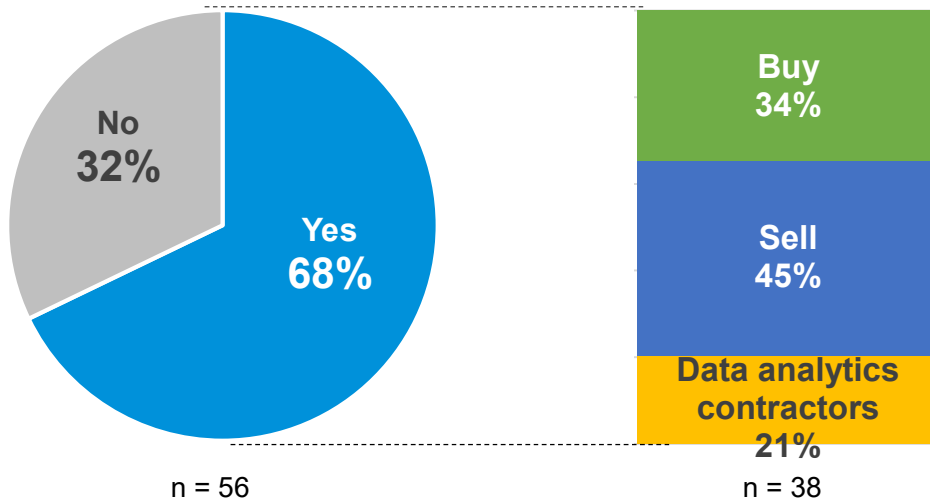
August 15, 2024 - October 15, 2024

Experience of Using Alternative Data

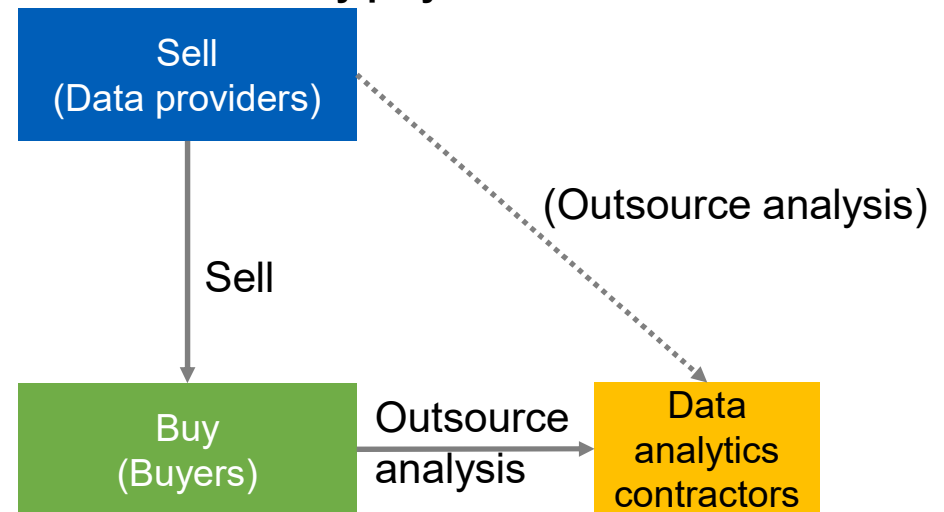
| | |
|----------------------------|----------------|
| Buyers | Data providers |
| Data analytics contractors | Non-users |

- To the question that asked whether they have experience of using alternative data, 68% answered “Yes.”
- The respondents who answered “Yes” have “Purchased” (34%), “Sold” (45%), or “Analyzed as a contractor” (21%).”
- We define respondents who have purchased alternative data as “buyers,” those who have sold data as “data providers” in the following questions.

■ Have you used alternative data?



■ Conceptual image of relationships between the key players

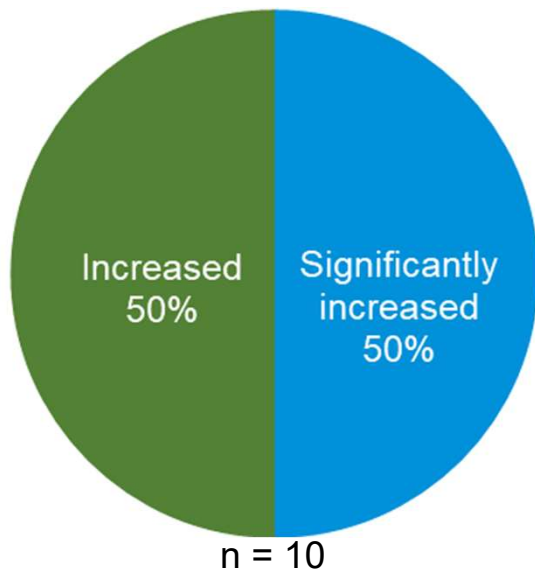


Changes in Expenditure on Alternative Data

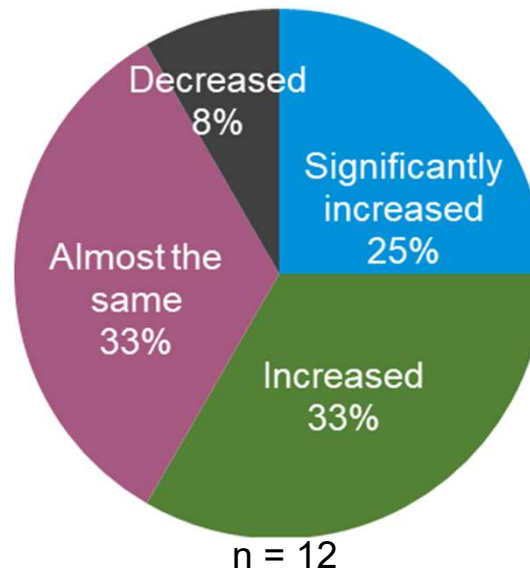
- We asked buyers how their expenditure on alternative data have changed compared to the previous year.
- 85% of respondents answered “Almost the same” and 15% “Increased”.
- We find a gradual slowdown in the rate of growth.

■ How has your spending on alterative data changed compared to previous year?

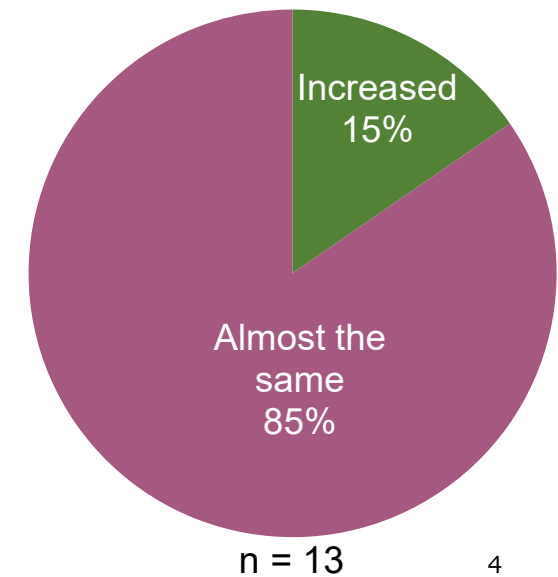
2022 survey



2023 survey



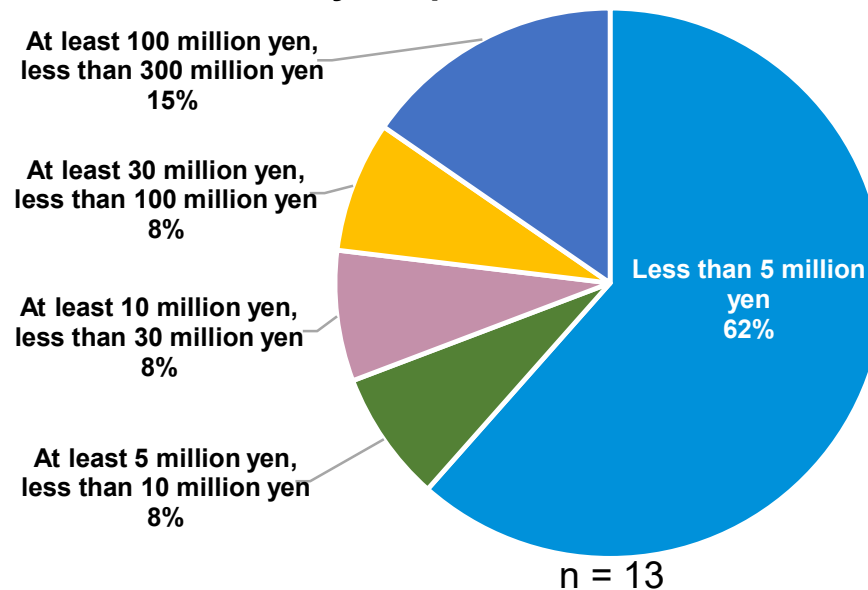
2024 survey



Ranges of Annual Expenditure

- We asked buyers about the range of their annual expenditure on alternative data.
- “Less than 5 million yen,” accounted for 62%, “At least 100 million yen, less than 300 million yen” for 15%, and the other answers were in the minority.
- Considering the expensive unit prices of alternative data, this result suggests that only a limited number of members use the data extensively and most buyers use the data to a limited extent.

■ How much do you spend on alternative data in a year?



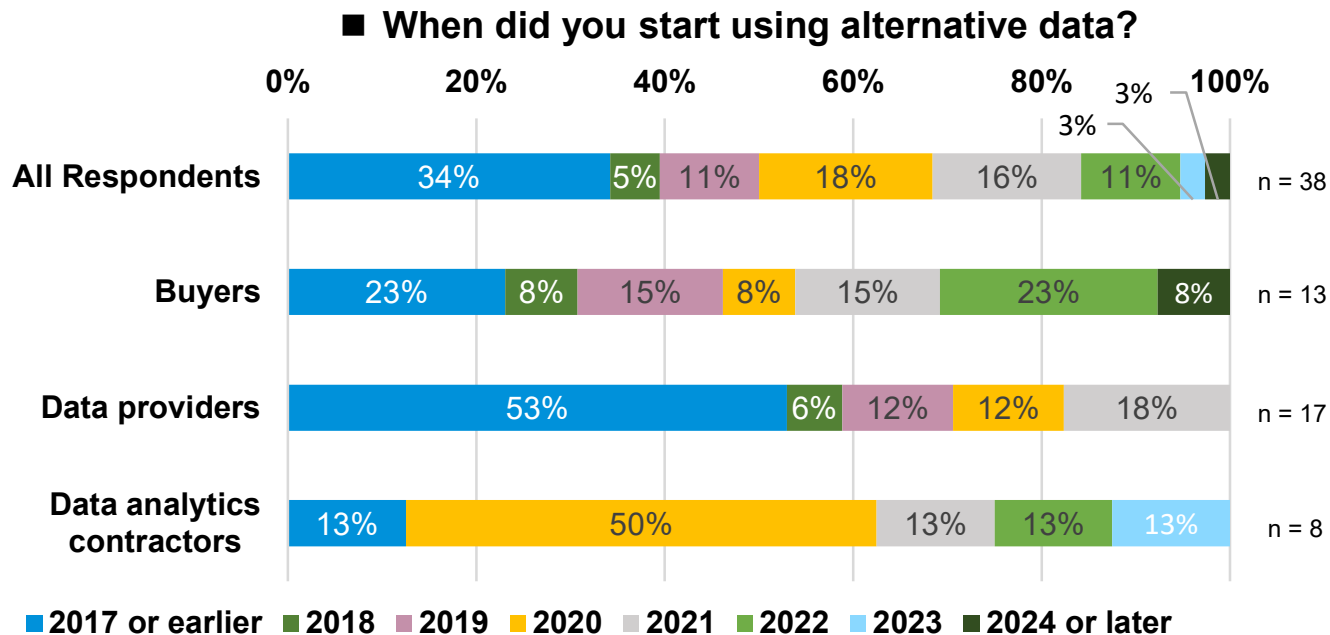
Start Data for Using Alternative Data

Buyers

Data providers

Data analytics contractors

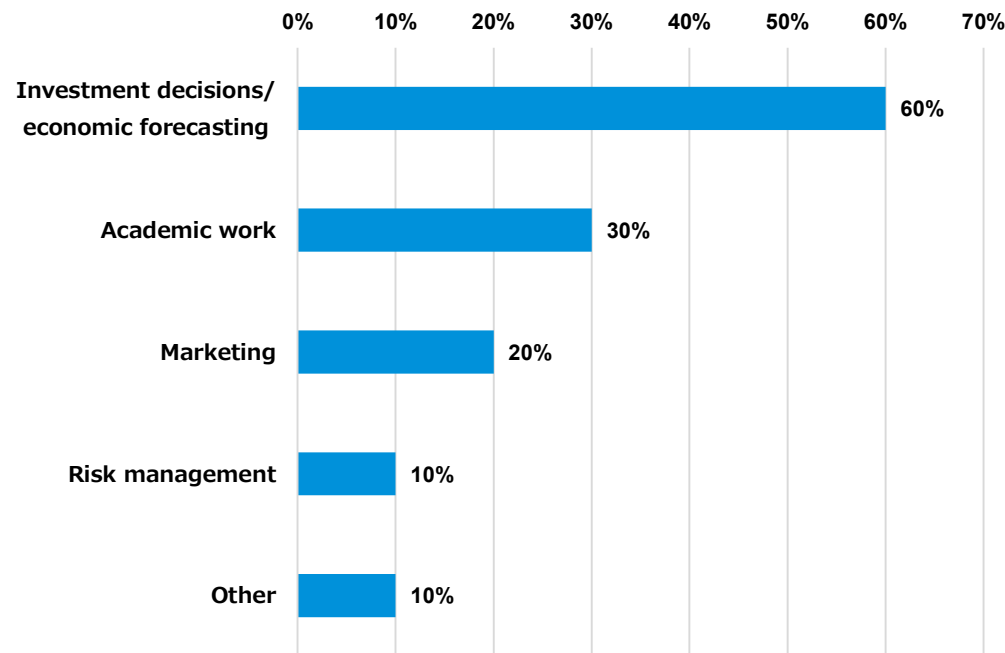
- To the question that asked about the timing they had started using alternative data, the most common answer among all respondents was “2017 or earlier.” However, those who started in 2020 or later also made up large percentages, indicating that demand for alternative data had grown during the pandemic.
- The breakdown by entities also follows the similar trend, but while data providers and data analytics contractors started using it early, there are also a reasonable number of buyers who started using data after 2022.



Purpose of Using Alternative Data

- We asked buyers the purpose of using alternative data.
- The majority answered “Investment decisions/economic forecasting” (60%), followed by “Academic work” (30%), “Marketing” (20%) and “Risk management” (10%).

■ What do you use alternative data for? (multiple choice question, up to 3 answers)

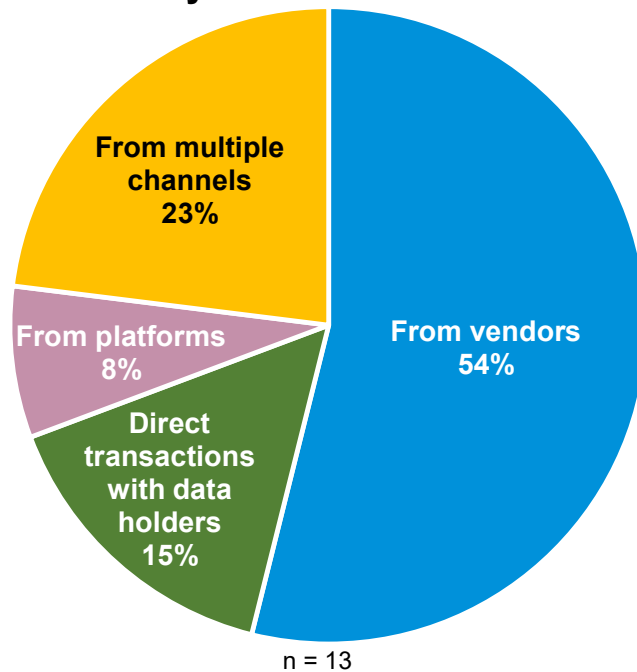


n = 10

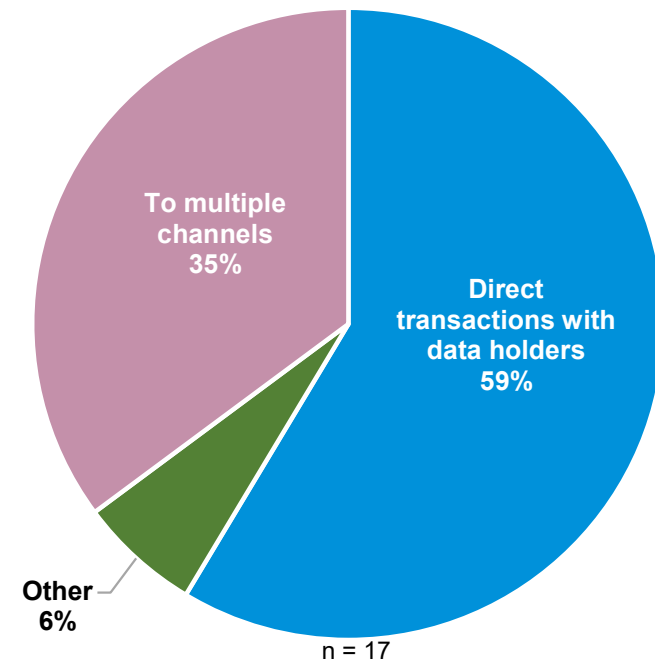
Channels of Purchasing/Selling Data

- We asked about the purchase/sales channels they use to buy/sell alternative data.
- The most common answer was “From vendors” among buyers, and was “Direct transactions” among data providers.

■ What purchase channel do you use to buy alternative data?



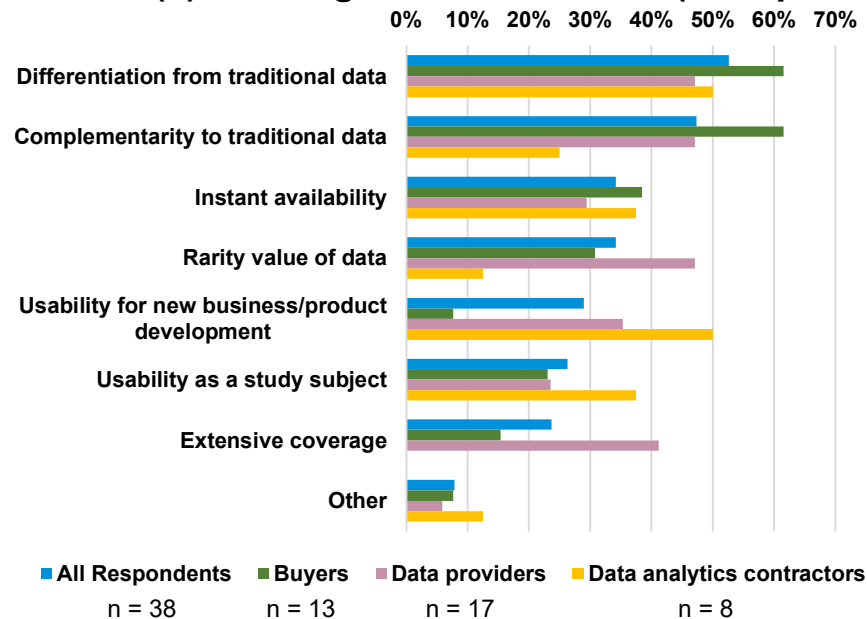
■ What sales channel do you use to sell alternative data?



Benefits of Using Alternative Data

- To the question about the benefits of using alternative data, 53% of all respondents selected “Differentiation from traditional data,” 47% answered “Complementarity to traditional data,” and 34% “Instant availability.”
- The result are not quite the same when responses were sorted by groups. The most common answers were “Differentiation from traditional data” and “Complementarity to traditional data” for buyers, whereas it is “Usability for new business/product development” for data providers.

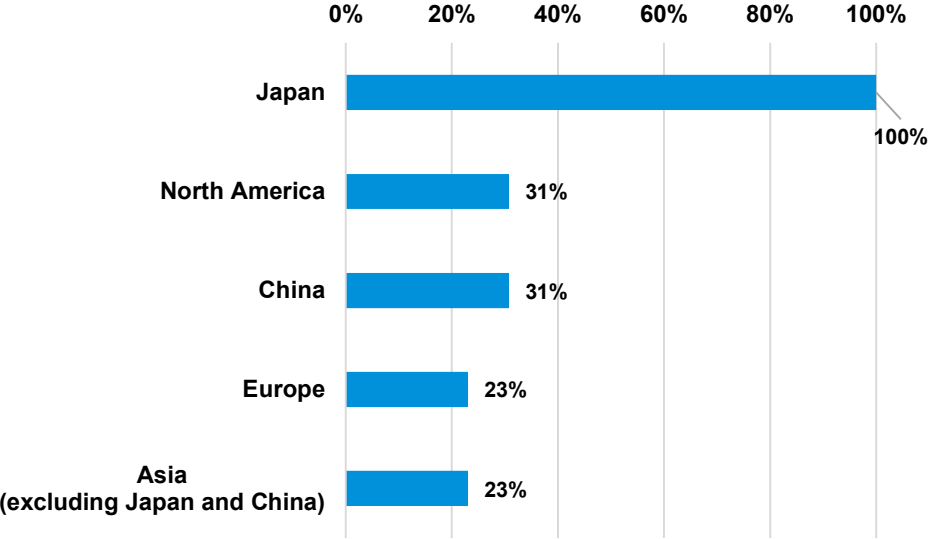
■ What is/are the benefit(s) of using alternative data? (multiple choice question, up to 3 answers)



Regions/Sectors of Alternative Data

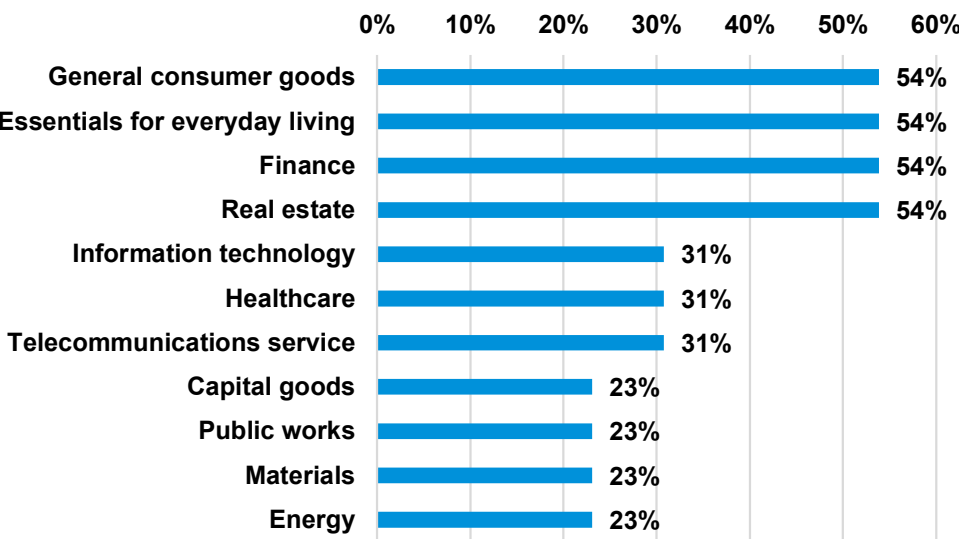
- We asked buyers the region(s) that the purchased alternative data covers. The most common answer was "Japan." For foreign regions, North America, China, Europe and other Asian countries were used to a certain extent.
- Regarding sector(s), the most common answer was "General consumer goods," followed by "Essentials for everyday living," "Finance," "Real estate," and "Information technology."

■ Regions covered by purchased alternative data (multiple choice question)



n = 13

■ Sectors covered by purchased alternative data (multiple choice question)



n = 13

Types of Alternative Data

Buyers

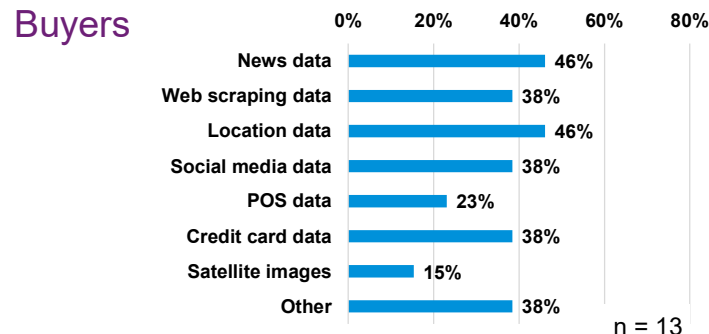
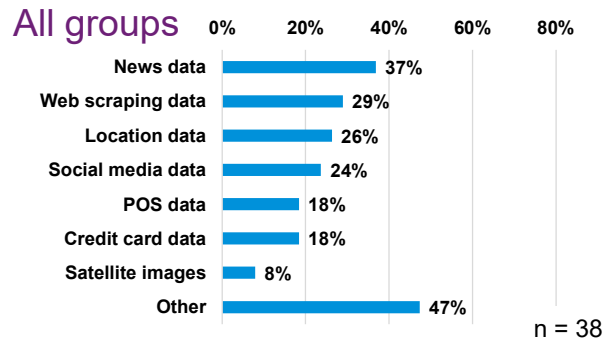
Data providers

Data analytics contractors

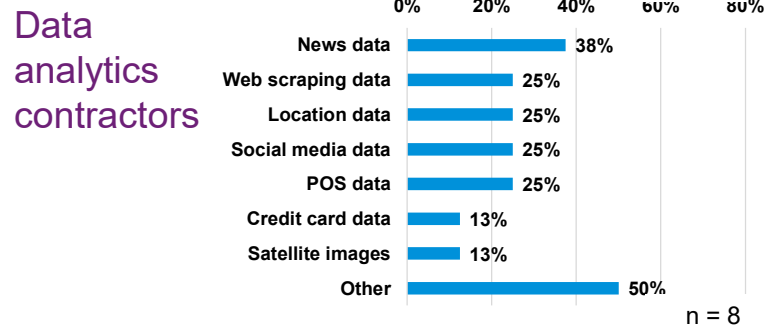
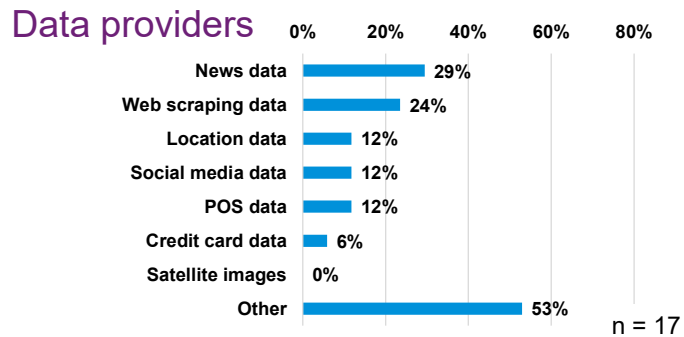
- In regard to the types of alternative data handled, most popular data for all respondents were “News data, followed by “Web scraping data” and “Location data”.
- For other data, it shows that a variety of data is being used, such as “Word of mouth data”, “Medical data” and “Television viewing data”.

■ Types of alternative data used (multiple choice question)

■ Other



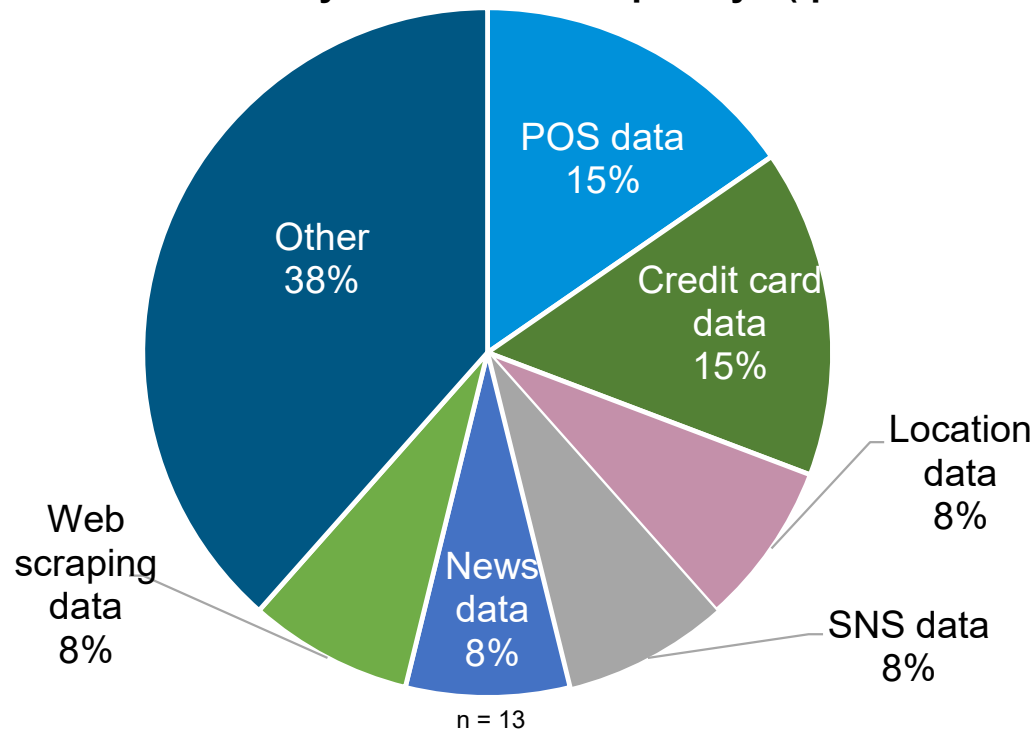
- Asset price data
- Word of mouth data
- Market related data
- Medical data
- Patent-related data
- Employee engagement data
- ESG data
- Access log data
- Television viewing data
- Financial data



Most Frequently Used Alternative Data

■ We asked buyers the type of alternative data they most frequently use. The most frequently used data are “POS data” (15%) and “Credit card data” (15%). Some answered “Location data” (8%), “SNS data” (8%), “News data” (8%), and “Web scraping data” (8%).

■ Which alternative data do you use most frequently? (question for buyers)



■ Other

- Asset price data
- Word of mouth data
- Market related data
- Smartphone application data
- Geographical data

Data Obtained By Web Scraping

Buyers

Data providers

Data analytics
contractors

- In 2024 survey we asked web scraping users about how they obtain and use the data. We found a wide variety of data is being scraped, such as publicly available information including “corporate disclosure information”, as well as “ X (formerly Twitter) data”.
- Regarding the usage of the data, the answers were “data provision” and “in-house use.”

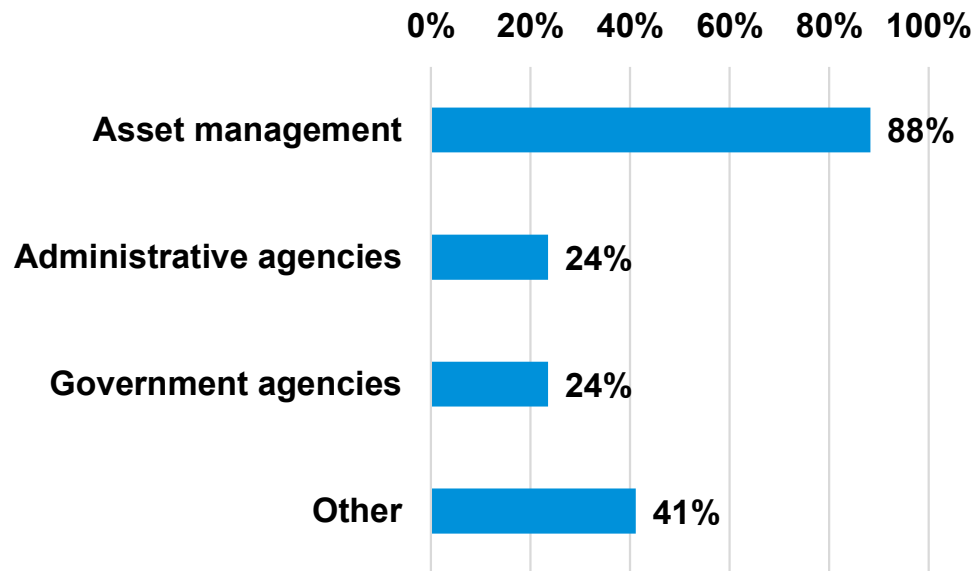
■ Obtainment and usage of web scraping (users of web scraping)

| |
|---|
| Corporate disclosure information. |
| Data related to crowdfunding (movement of fund procurement status, etc.). |
| Data related to job advertisements. |
| X (formerly Twitter) data. |
| The status of shared office openings by our company and other companies. |
| Disclosure information from companies. |
| Various announcements on the website. |
| Using AI and algorithms, we collect data mechanically, cleanse, merge and sort names, and provide the data or provide it in the form of scores. |
| Scraping public websites to obtain licensed content via third-party data providers and using it for in-house data. |

Customers of Alternative Data Providers

- We asked data providers to whom they sell alternative data.
- The most common answer was “Asset management companies” (88%), followed by “Administrative agencies” (24%) and “Government agencies” (24%).

■ **Who buys the alternative data you offer?**
 (multiple choice question)



n = 17

■ **Other**

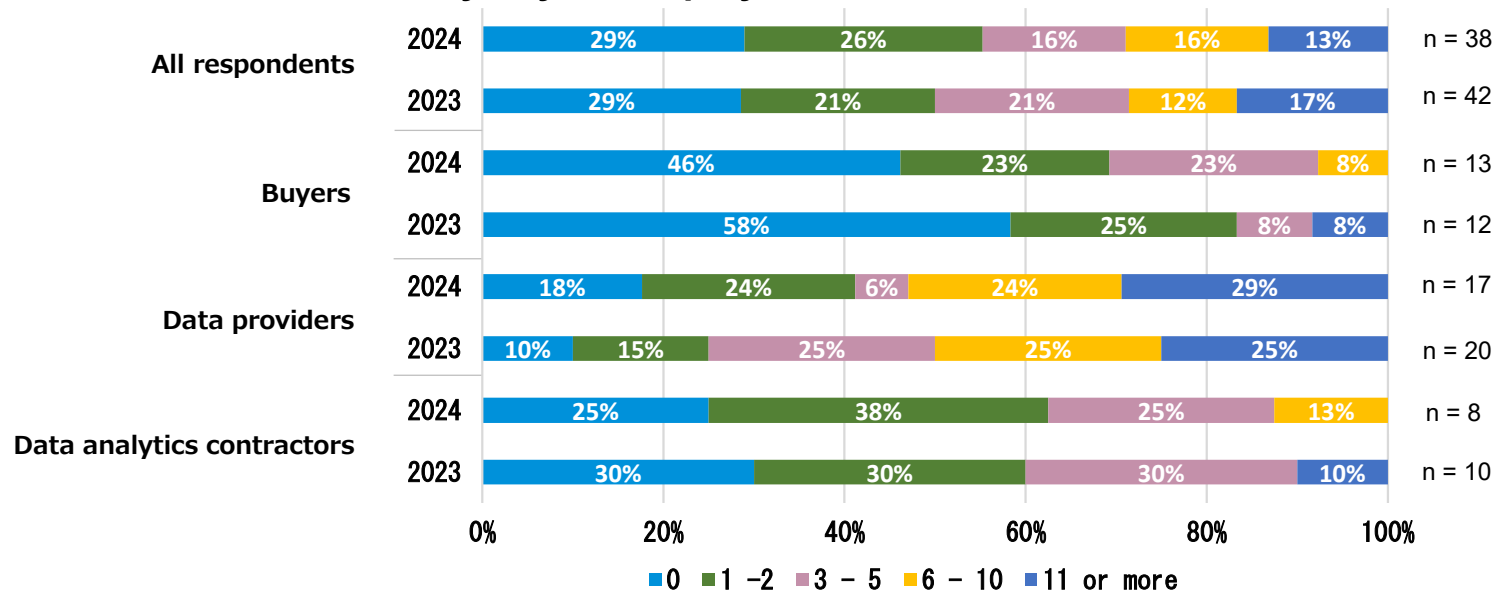
- Universities
- Businesses
- Information vendors

Employees Dedicated to Data Management

Buyers
Data providers
Data analytics contractors

- We asked the number of employees who are dedicated to alternative data management. We find that there were significant differences in trends between the groups.
- While many data providers and data analysis contractors dedicated several staffs, almost a half did not dedicate any staff among buyers.
- However, compared to 2023 survey, the number of dedicated staffs in buyers has increased, indicating that their structures of alternative data analysis are improving.

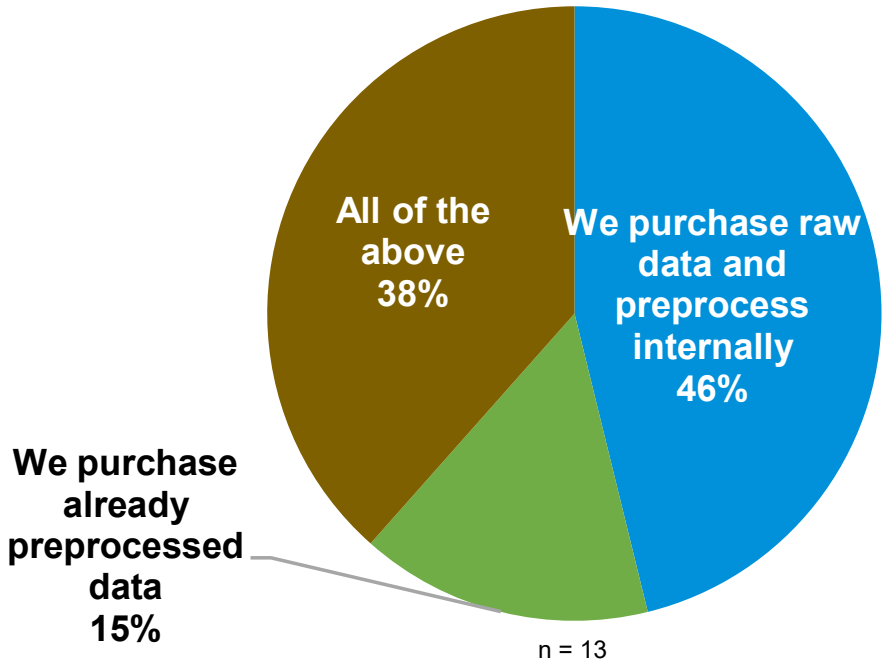
■ How many of your employees are dedicated to alternative data management?



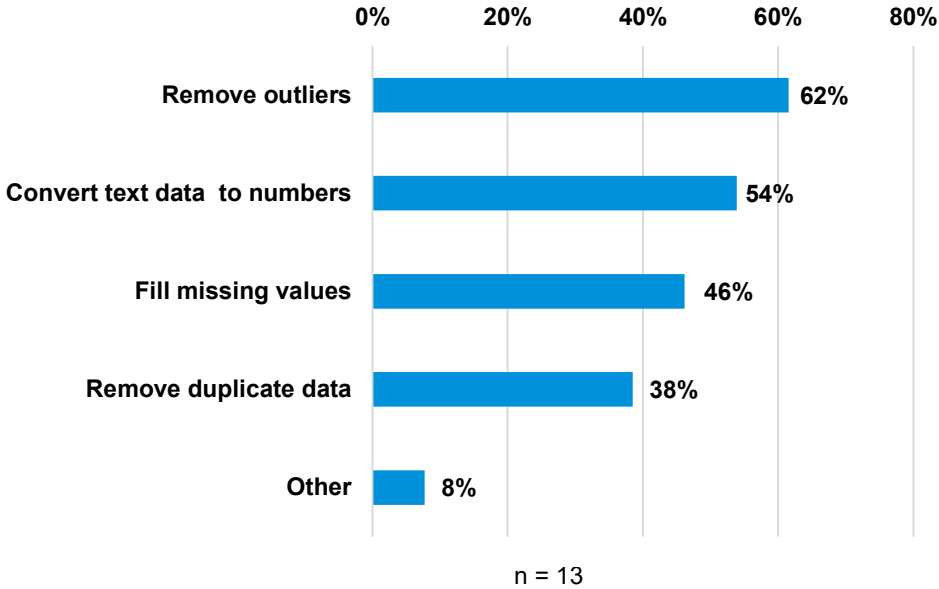
Preprocessing of Alternative Data

- We asked buyers the preprocessing of alternative data. The results show that the majority of respondents preprocess alternative data either by themselves. The increasing number of dedicated staffs among buyers may make it possible to preprocess internally.
- To the question that asked the part of preprocessing, the majority answered “Remove outliers” (62%), “Convert text data to numbers” (54%), and “Fill missing values” (46%).

■ How do you preprocess alternative data you purchase?



■ Which part(s) of data do you preprocess? (multiple choice question)



Regulations That Need Improvement

Buyers

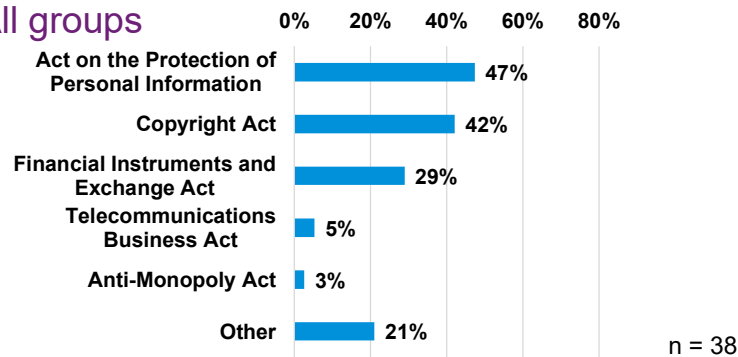
Data providers

Data analytics
contractors

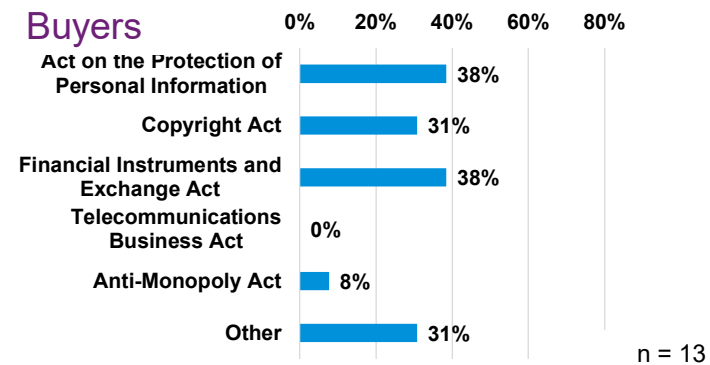
- To the question that asked about regulations that need improvement to promote the use of alternative data, the common answers from all respondents were the “Act on the Protection of Personal Information” (47%), “Copyright Act” (42%), and “Financial Instruments and Exchange Act” (29%).
- Showing detailed difference by groups, data providers and data analysis contractors consider that “Act on the Protection of Personal Information” and “Copyright Act” were in particular needed to be improved.

■ Laws that need improvement to promote the use of alternative data (multiple choice question, up to 3 answers)

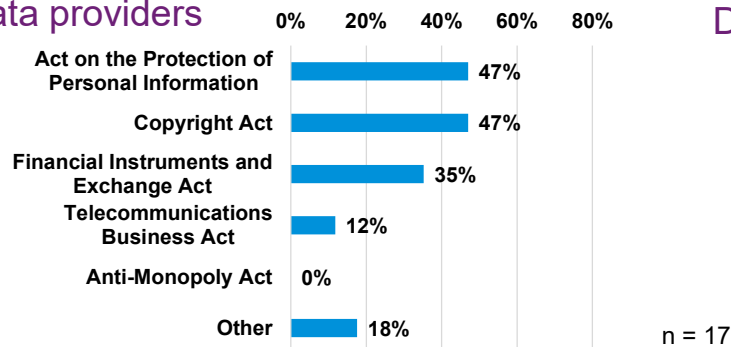
All groups



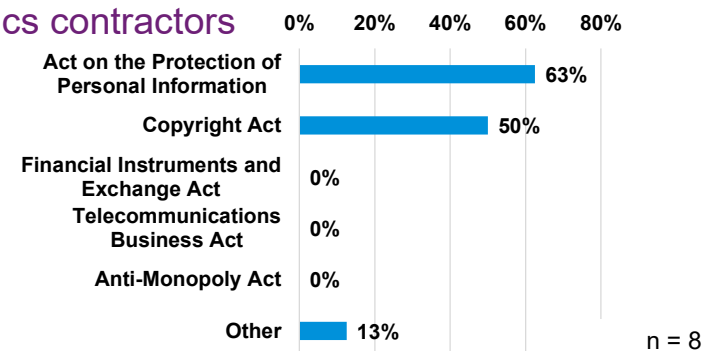
Buyers



Data providers



Data analytics contractors



























Regulations That Need Improvement (sorted by Type of Data)

Buyers
Data analytics
contractors

Data providers

- We then cross-tabulated the responses, sorting by the type of alternative data they handle and the regulations that need improvement to promote the use of alternative data.
- “Act on the Protection of Personal Information” was all in all believed to need improvement, regardless of the type of data handled. For the “Copyright Act”, answers varied according to the type of data used. It was commonly believed to need improvement by users of “POS data and “News data”, whereas the proportion were lower among users of “Credit card data and “Location data”.
- As for the “Anti-Monopoly Act” and the “Telecommunications Business Act”, smaller proportions of users answered that these laws need improvement compared to the other laws, regardless of type of data handled.

■ Laws That Need Improvement (Sorted by Type of Data Used, multiple choice question)

| | Act on the Protection of Personal Information | Copyright Act | Financial Instruments and Exchange Act | Anti-Monopoly Act | Telecommunications Business Act | n |
|----------------------|---|---|---|---|---|----|
| POS |  43% |  57% |  29% |  14% | 0% | 7 |
| Credit card |  43% |  14% |  14% | 0% | 0% | 7 |
| Location information |  30% |  10% |  20% | 0% |  10% | 10 |
| Satellite images |  33% |  33% |  33% | 0% | 0% | 3 |
| Social media |  56% |  33% |  44% | 0% | 0% | 9 |
| Web scraping |  56% |  33% |  44% | 0% | 0% | 11 |
| News |  43% |  57% |  43% |  7% | 0% | 14 |

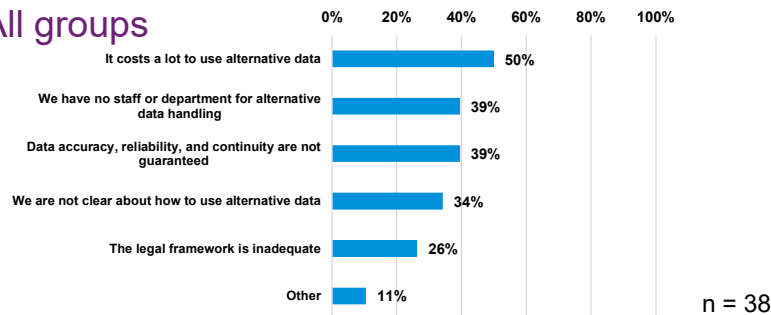
Issues Involved in the Use of Data

Buyers
 Data providers
 Data analytics contractors

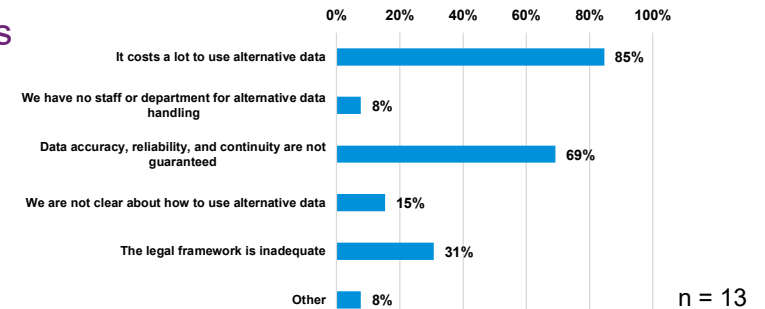
- We asked about issues involved in the use of alternative data. For all respondents, common answers were “It costs a lot to use alternative data,” followed by “We have no staff for alternative data handling,” “Data accuracy, reliability, and continuity are not guaranteed”.
- For buyers, the most common answer was the issue of costs. For data providers, a large proportion selected the issues of staff and of uncertainty about using alternative data. For data analytics contractors, the most common answer was the issue of staff, followed by issues of data accuracy, reliability and continuity.

■ Issues involved in the use of alternative data (multiple choice question, up to 3 answers)

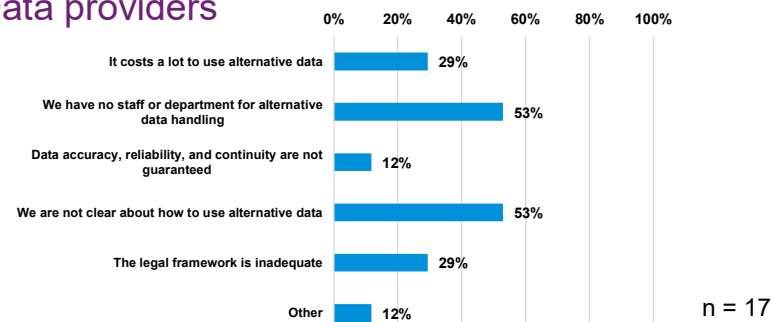
All groups



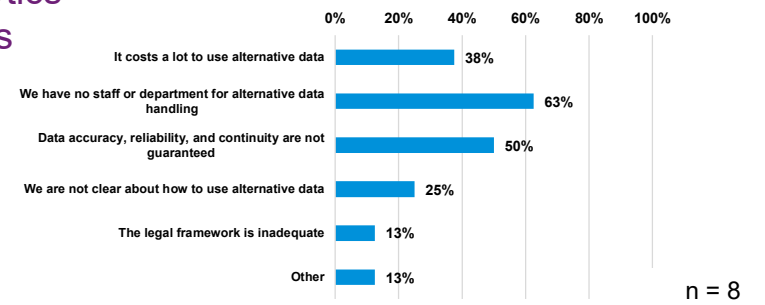
Buyers



Data providers



Data analytics contractors



Issues (Buyers)

■ Issues involved in the use of alternative data (provided by buyers, free-form answers)

Insufficient data accumulation makes it difficult to understand the habits of statistics and to perform seasonal adjustments.

It is difficult to use as a main data source because of its poor accuracy.

I have the impression that alternative data is being sold at high prices without a clear idea of its specific use cases.

The data obtained is only valid under specific conditions, so it is not very versatile. There is a possibility of reaching the wrong decision.

More expensive than traditional data.

There are some data that are prohibited from being used in generative AI, and you have to pay a fee for each project.

Because it is not possible to try using alternative data when asked about cost-effectiveness within the company.

Many of the data sets cost over 5 million yen, so it is difficult to introduce them unless there is a good chance of a return on investment.

Since many of the alternative data are new, there are many cases where there is not enough time to perform backtesting.

Issues (Data Providers)

■ Issues involved in the use of alternative data (provided by data providers, free-form answers)

Because this data is not included in the previous workflow, it is difficult for customers to imagine how to use it.

Data of a different type to existing data will require individual handling.

It takes time to train people who can handle data and to disseminate data itself.

Because data is expensive, it is often the case that it will not lead to a purchase unless the benefits are shown to be worth more than the price or a discount is offered.

Many Japanese companies (particularly general business companies) lack the human resources to be able to use alternative data.

It needs to be made easier to handle.

In particular, domestic companies' understanding of alternative data is still not very advanced.

I have the feeling that there is not enough of a system or budget in place to carry out the analysis.

Because the data is so advanced, there are often cases where people don't know how to use it.



Issues (Data Analytics Contractors)

Data analytics
contractors

■ Issues involved in the use of alternative data (provided by data analytics contractors, free-form answers)

As there is no established technical system for handling unstructured data, considerable ingenuity is required for each type of data.

If the premise of data collection is not known, questions remain about the reliability and accuracy of the data.

It is difficult to recruit staff to handle data.

Proof of Concept (PoC) is required to verify the accuracy of the data, which increases the man-hours required.

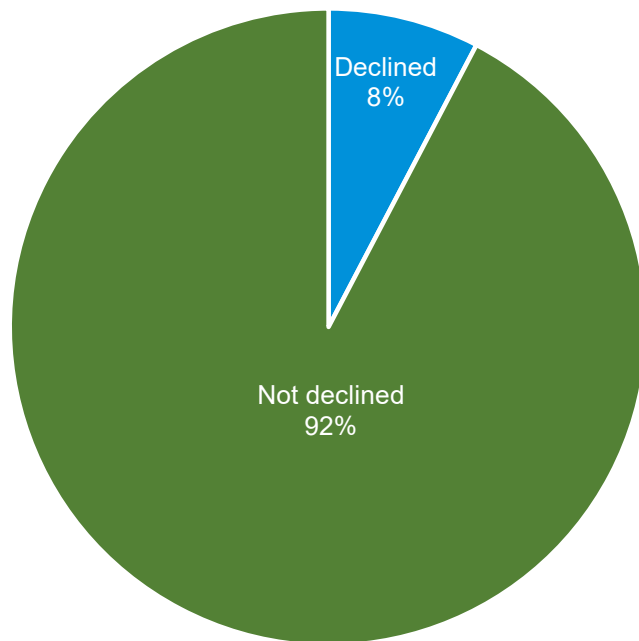
As the trends in data change one after another, it is extremely burdensome in terms of business to keep track of each one of them.

There is a shortage of people with skills in data processing and analysis.

Need for Alternative Data

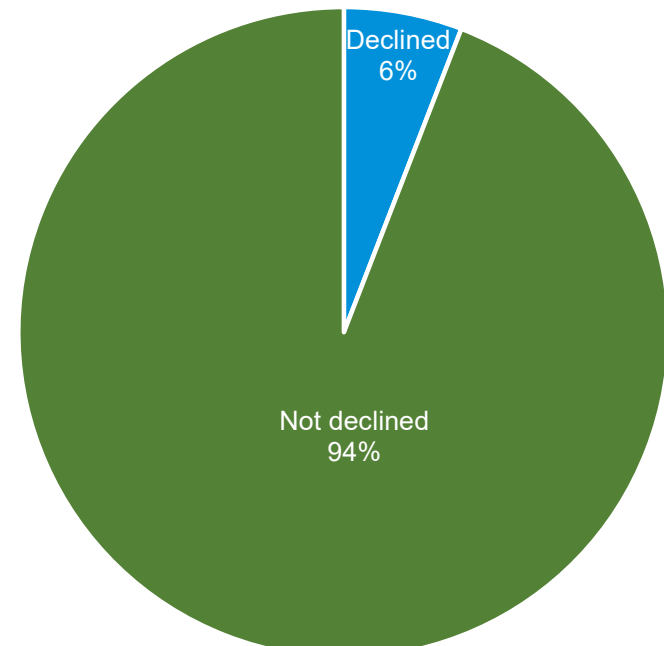
■ We asked buyers about their needs for alternative data, and data providers about their customers' needs. The result indicates that there was still a strong demand for alternative data, despite the decline in specific demand seen during the COVID-19 period.

■ Is the need for alternative data decreasing?



n = 13

■ Are customer needs for alternative data declining?



n = 17

Areas where demand has not (has) declined

■ Areas where demand is not declined

Asset management, investment banking, insurance, and business companies.

Consumer behaviour and marketing, sustainability-related, healthcare.

Text data from newspapers, social networking sites, etc.

Wages and employment-related fields.

Data that can be expected to be used in generative AI and in the macro market.

Marketing and insurance (pre-symptomatic) domains.

Financial Instruments and Exchange Data.

Non-financial information/

Sales field.

Insurance premium rate.

Company Information.

■ Area where demand is declining

Costly data.

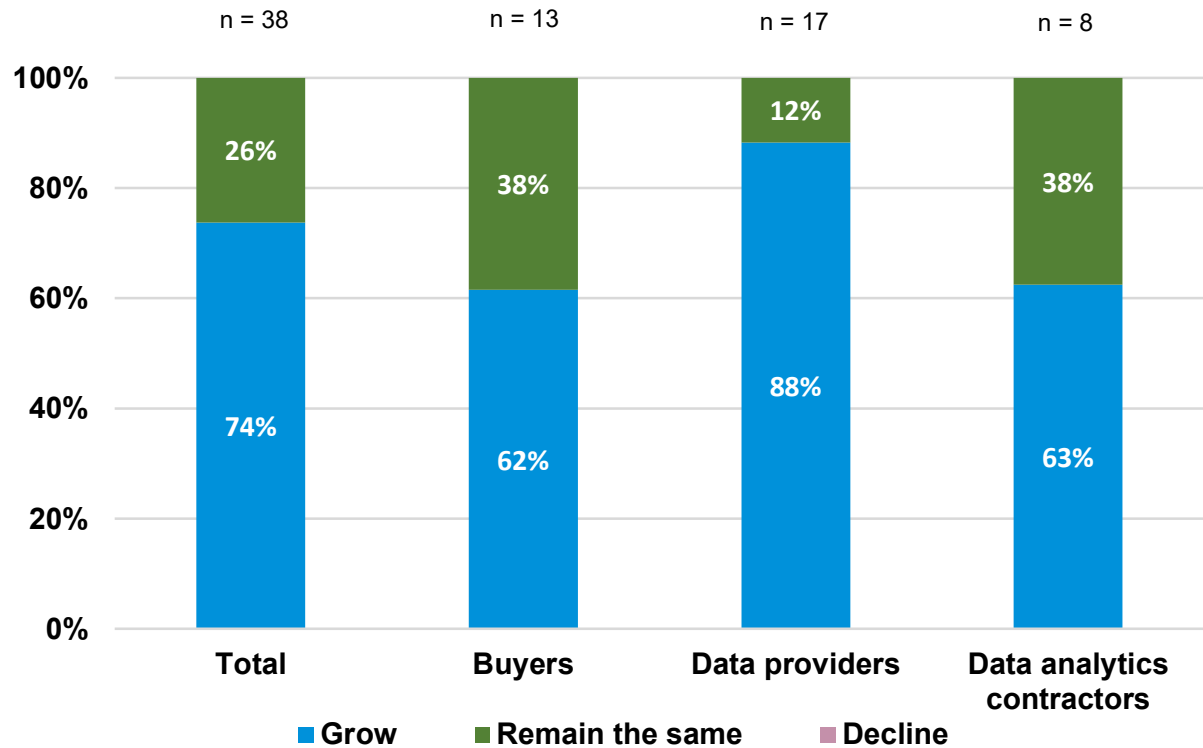
Asset Management Field.

Outlook for Alternative Data

Buyers
 Data providers
 Data analytics contractors

■ To the question about the outlook for the alternative data market (over the next three years, starting from this year), the majority of respondents answered that “the market will grow.”

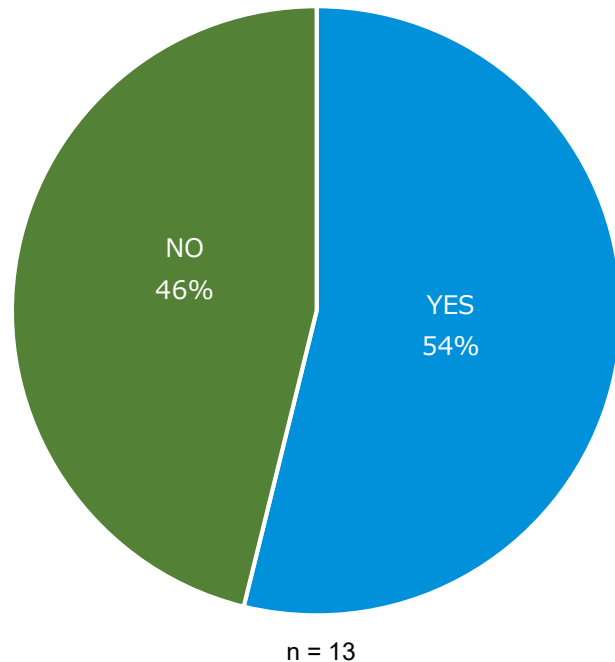
■ Do you think the alternative data market will grow in the future (over the next three years, starting from this year)?



Using Alternative Data for AI

■ In 2024 survey, we asked buyers about the experience of using AI in alternative data analysis, 54% of buyers answered they had such experience. Regarding specific methods and data usage, the answers included “translation and summarization of text data”, “use in quantitative funds”, indicating that it is being used in a wide range of fields.

■ **Have you ever used AI in business or research using for alternative data?**



■ **What are the specific ways of using it and what data is used?**

- Investment decisions and the streamlining and sophistication of the process
- Translation and summarization of text data
- Use in quantitative funds
- Analysis of consumer behavior
- Model development

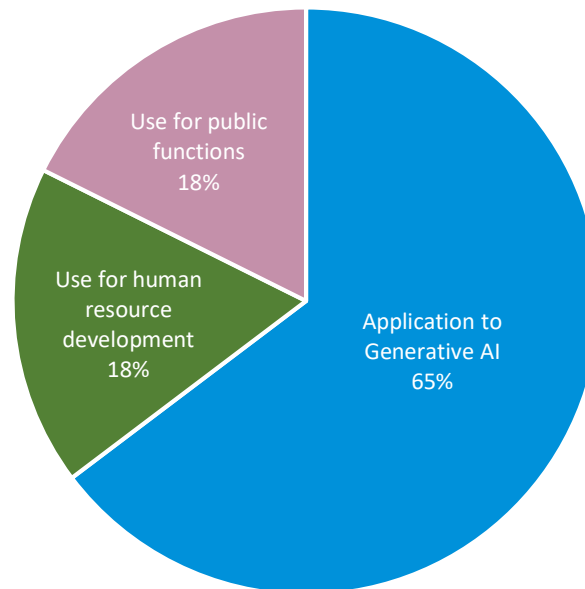
■ **Is there a reason you're not using it?**

- We are using AI, but it is not linked to alternative data.
- I can't imagine a specific example of how to use it.

Future Use Cases of Alternative Data

- In 2024 survey, we asked data providers which alternative data use cases that they expected to grow in the future. The majority of response was “Application to Generative AI” (65%), followed by “Use for human resource development” (18%), and “Use for public function” (18%).

- **What kinds of alternative data use cases do you think will grow in the future?**



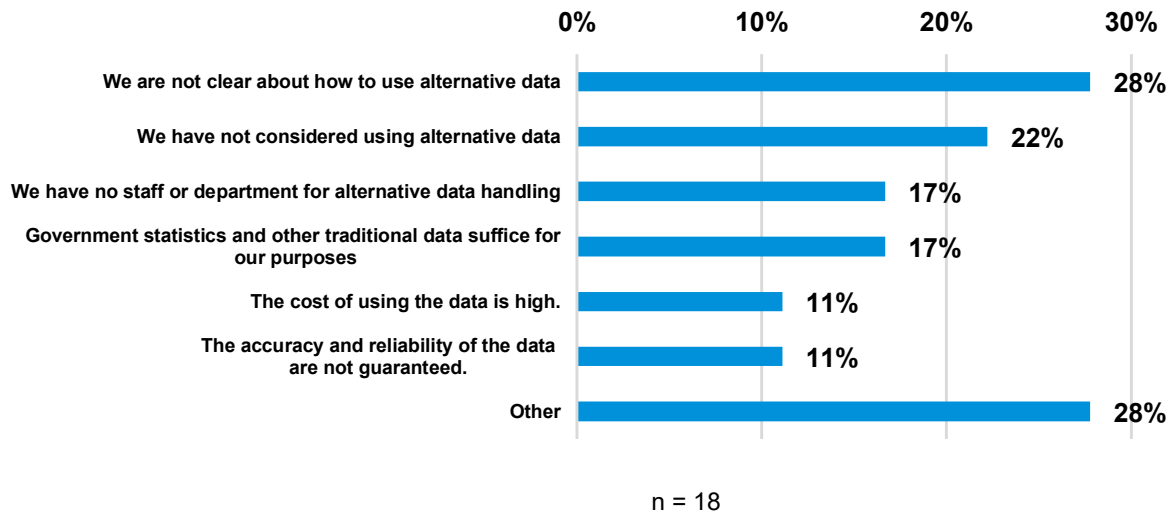
n = 17

Reason for Not Using Alternative Data

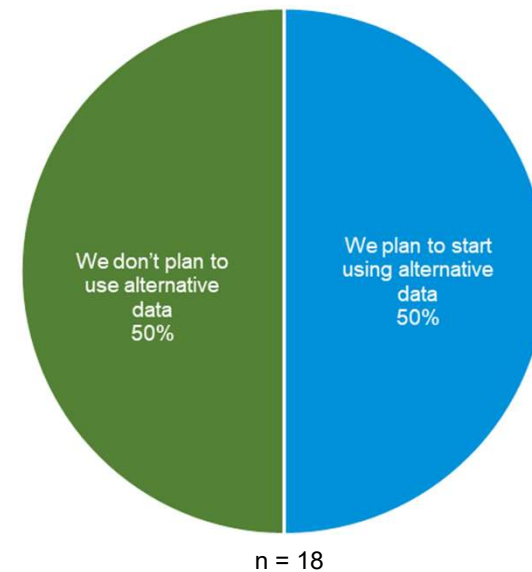
Non-users

- We asked non-users of alternative data why they did not use it. The most selected reason was “We are not clear about how to use alternative data” (28%), followed by “We have not considered using alternative data” (22%), “We have no staff or department for alternative data handling” (17%), and “Government statistics and other traditional data suffice for our purposes” (17%).
- However, half of these non-users intend to start using alternative data in the future.

Reasons for not using alternative data (multiple choice question, up to 3 answers)



Plan to use alternative data sometime in the future (the next three years, starting from this year)





Conclusions

Through the third year's FACTBOOK survey, we analyzed the quantitative evaluation of the use of alternative data in Japan. The results are summarized below:

- First, there is still a high level of demand for alternative data. Although the pace of increase in users' expenditure on the data was slowing, more than 90% responded that "demand for alternative data has not decreased," and expected that data market will continue to expand.
- Second, the number of experts who deal with alternative data is expanding. Even in buyers who claimed the shortage of experts, the number of dedicated staffs for alternative data has increased, which allowed them to purchase alternative data and make pre-process internally.
- Third, there are still various issues needed to be addressed to expand the use of alternative data. Cost of usage, difficulty in recruiting and training staffs to handle data, and concern about inadequacy of legal regulations are challenges to be solved for further use of the data.

We hope to continue this survey regularly to analyze trends in the use of alternative data, if members of the JADAA would agree to participate it. For the continuation of the survey, we would appreciate to receive your feedback and opinions.



JADAA

オルタナティブデータ推進協議会

Commissioned by: The Japan Alternative Data Accelerator Association

Commissioned: Sompo Institute Plus Inc.