



JADAA

オルタナティブデータ推進協議会

Alternative Data Factbook (Digest)

January, 2024

Japan Alternative Data Accelerator Association

Sompo Institute Plus Inc.

Survey Overview

Issue to explore

- Overseas countries are building up knowledge and insights on alternative data, and their research and studies on data use are also progressing.
- In Japan, there is not enough research on the use of alternative data, such as "who uses alternative 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 the data.

About this survey

- To gain a clear understanding about the use of alternative data and issues involved in it in Japan, we conducted a survey of members of the Japan Alternative Data Accelerator Association (JADAA).
- We compiled this Factbook from answers provided by 54 respondents.

■ Survey respondents

Business	Number of Respondents	Ratio
Asset management	3	6%
Finance and insurance (excluding asset management)	7	13%
Manufacturing	3	6%
Think tank	8	15%
Education	1	2%
Information technology/System development	24	44%
Other	8	15%
Total	54	100%

■ Survey data collection period

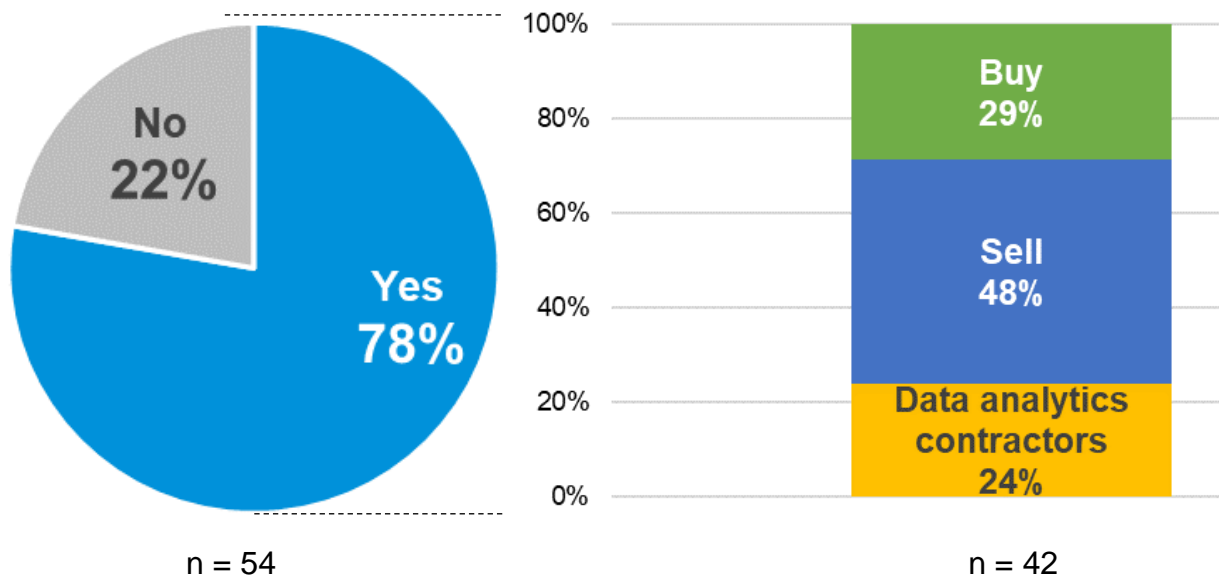
August 7, 2023 - September 15, 2023

Experience of Using Alternative Data

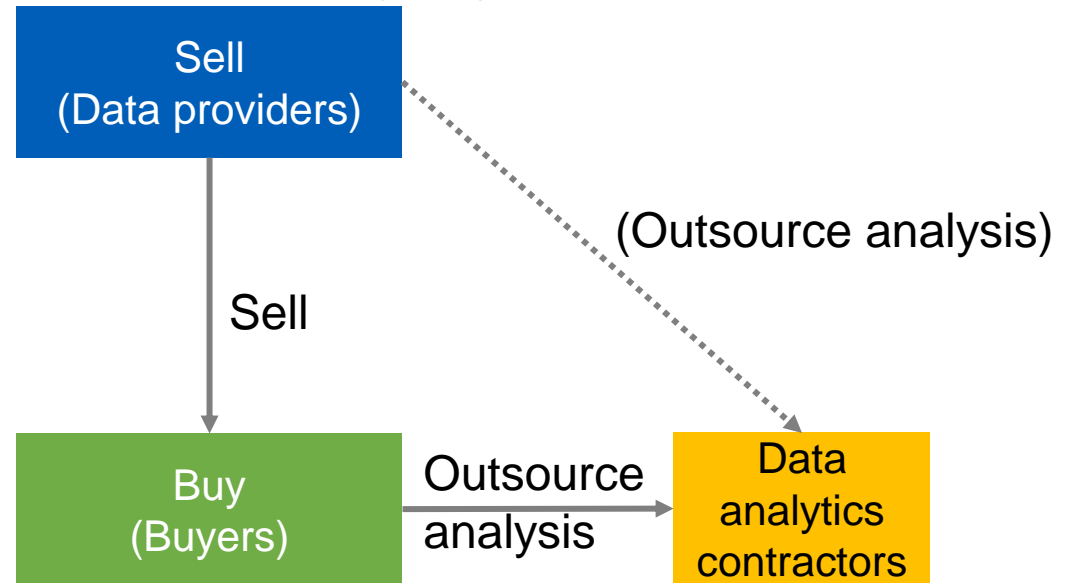
Buyers	Data providers
Data analytics contractors	Non-users

- To the question that asked whether they have experience of using alternative data, 78% answered “Yes.”
- The respondents who answered “Yes” have “Purchased” (29%), “Sold” (48%), or “Analyzed as a contractor” (24%).”
- 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 Spending on Alternative data

- We asked buyers how their spending on alternative data have changed compared to the previous year.
- 58% answered “Significantly increased” or “Increased,” 33% “Almost the same,” and 8% “Decreased.”
- Compared to the survey in 2022, although the pace of increase slowed down, increasing trend itself was continuing.

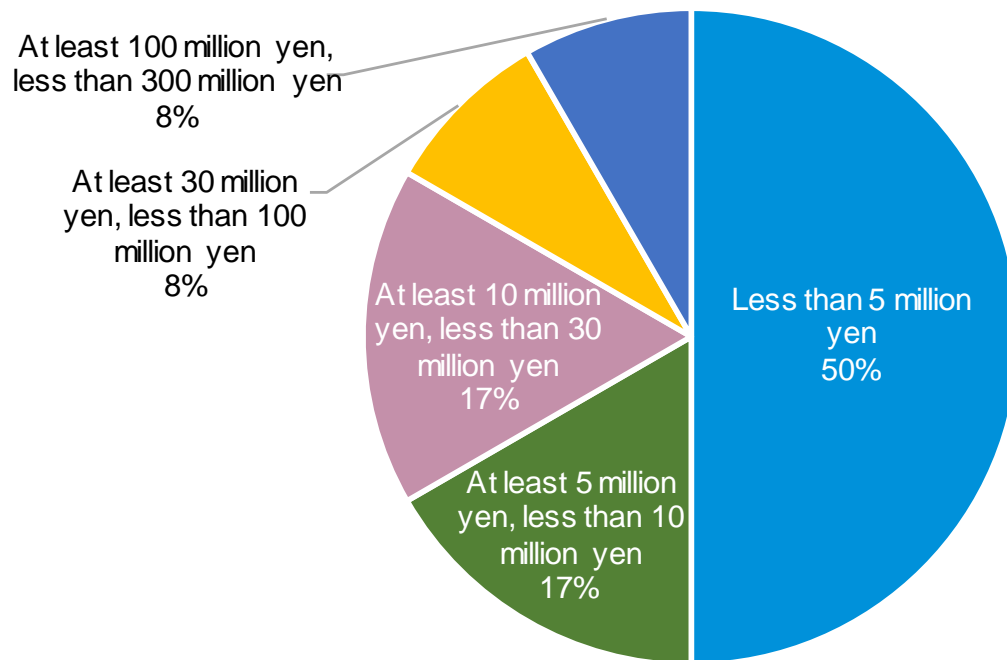
■ How has your spending on alternative data changed compared to previous year?



Ranges of Annual Spending

- We asked buyers about the range of their annual spending on alternative data.
- 50% answered “Less than 5 million yen,” and 17% answered “At least 5 million yen, less than 10 million yen” and “At least 10 million yen, less than 30 million yen” each. Respondents who answered that they spent “30 million yen or more” are less than 20%.
- Considering the unit prices of alternative data, this result suggests that only a limited number of companies 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?



n = 12

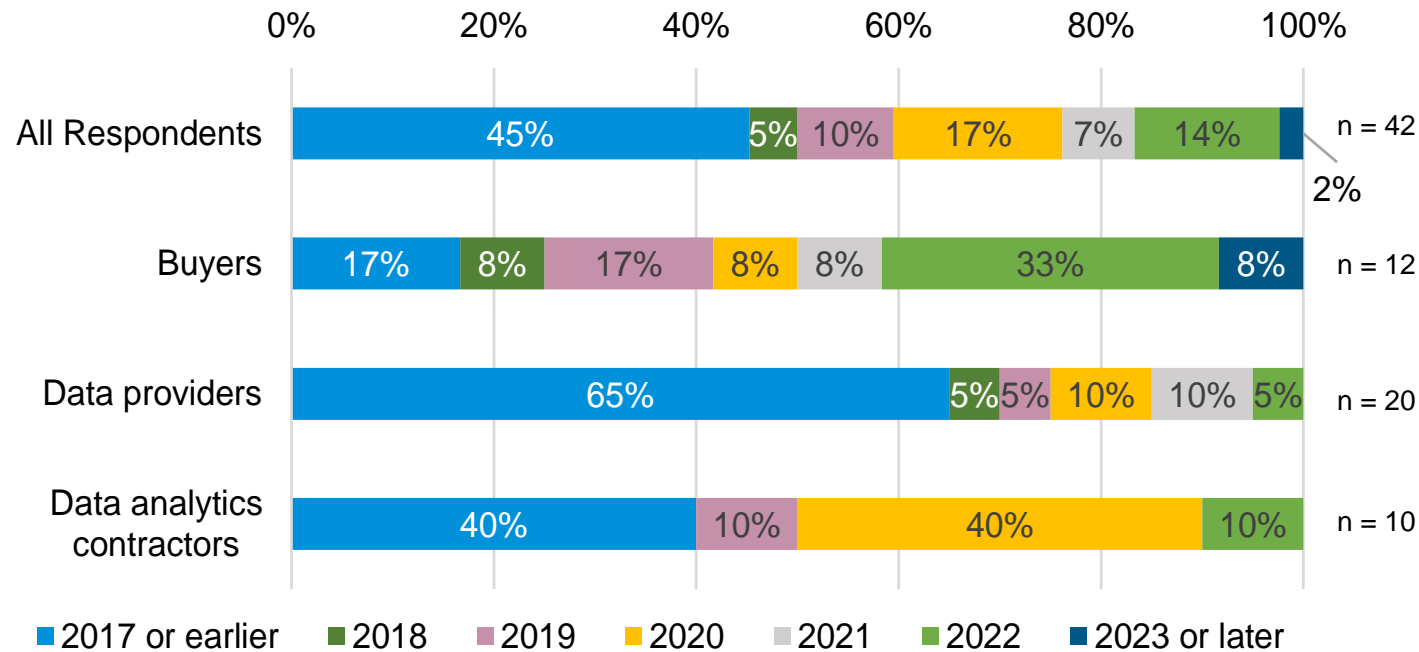
Start Data of Using Alternative Data

Buyers
 Data analytics contractors

Data providers

- 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 2018 or later also made up large percentages, indicating that demand for alternative data grew over the pandemic years.
- The responses sorted by player are also in line with the overall trends, whereas data providers began to use the data earlier than buyers.

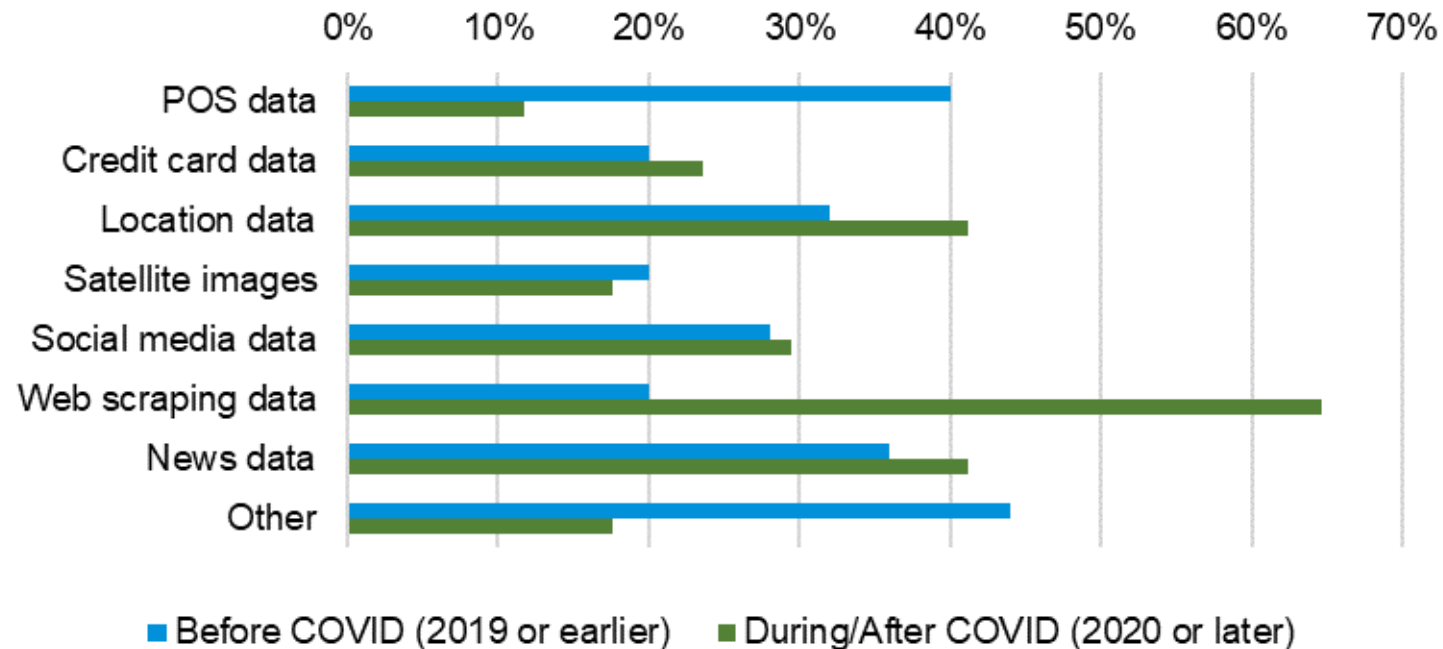
■ When did you start using alternative data?



Types of Data and Start to Use

- We next conducted cross-tabulation of responses from buyers, data providers, and data analytics contractors, based on the timing they began to use alternative data and types of data they use.
- Respondents who have used alternative data before the pandemic turned out to use “POS data” and “Location data” than other groups of respondents. On the other hand, respondents who began to use alternative data after the pandemic commonly use “Web scraping data” and “Location data”. These results revealed that the types of alternative data used vary according to when the use began.

■ Types of data used sorted by when the use began (multiple choice question)



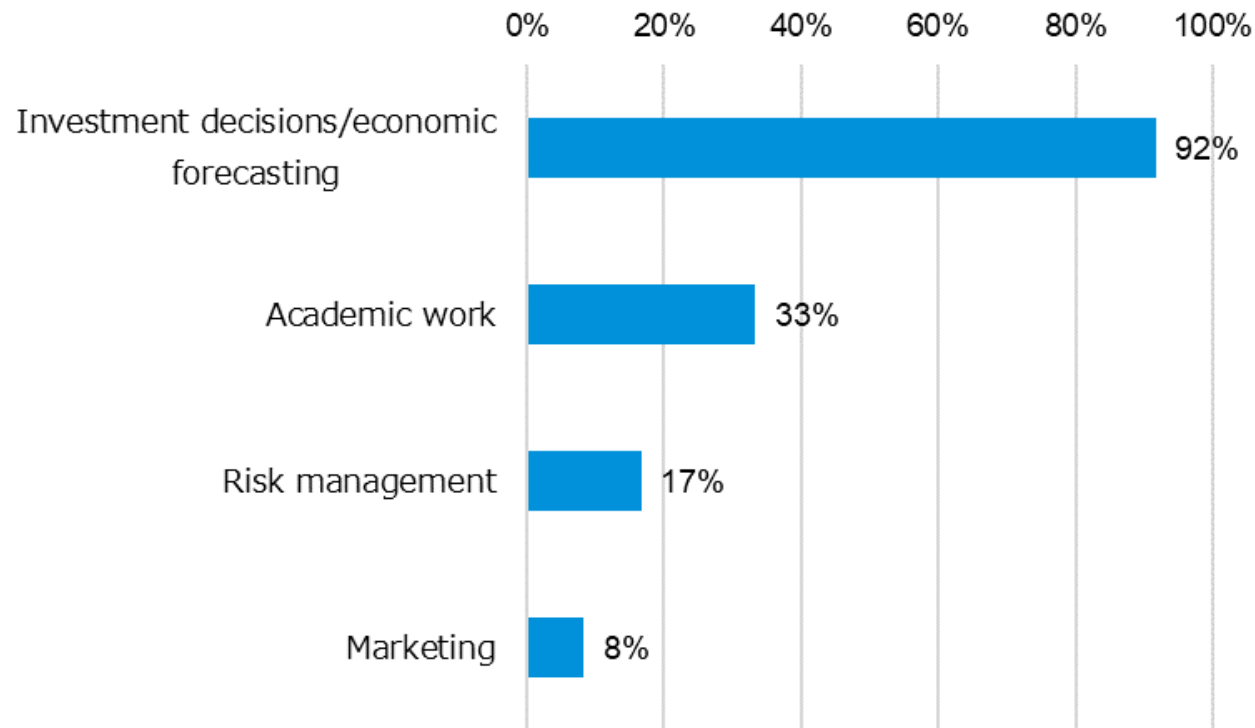
n = 25

n = 17

Purpose of Using Alternative Data

- We asked buyers the purpose of using alternative data for.
- The majority answered “Investment decisions/economic forecasting” (92%), followed by “Academic work” (33%), “Risk management” (17%), and “Marketing” (8%).

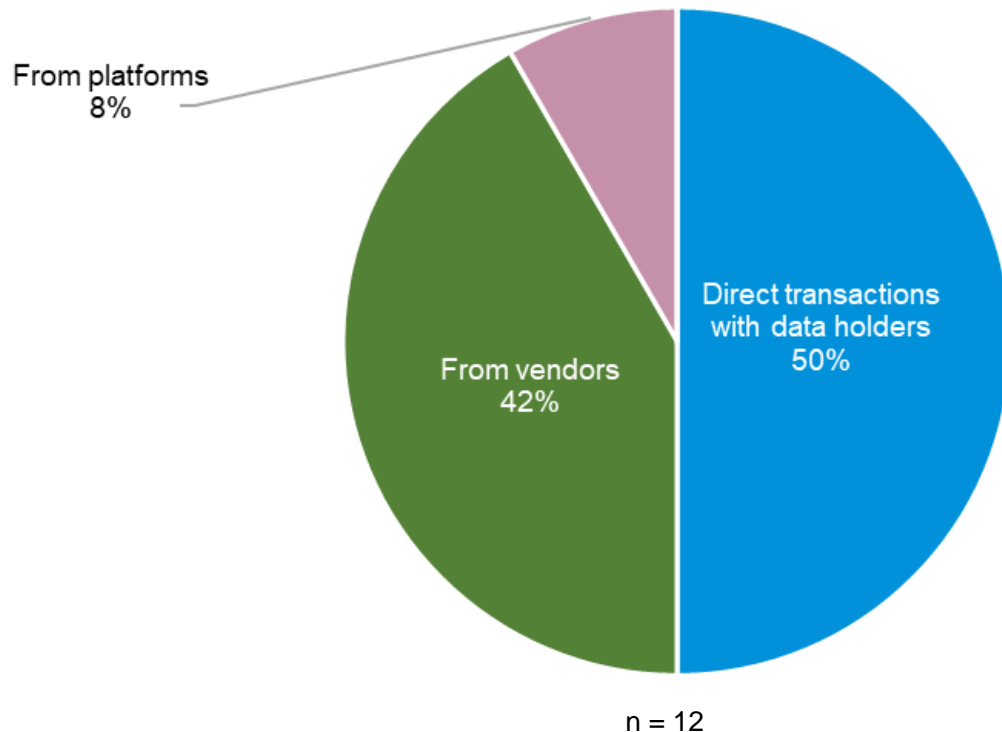
■ What do you use alternative data for? (multiple choice question, up to 3 answers)



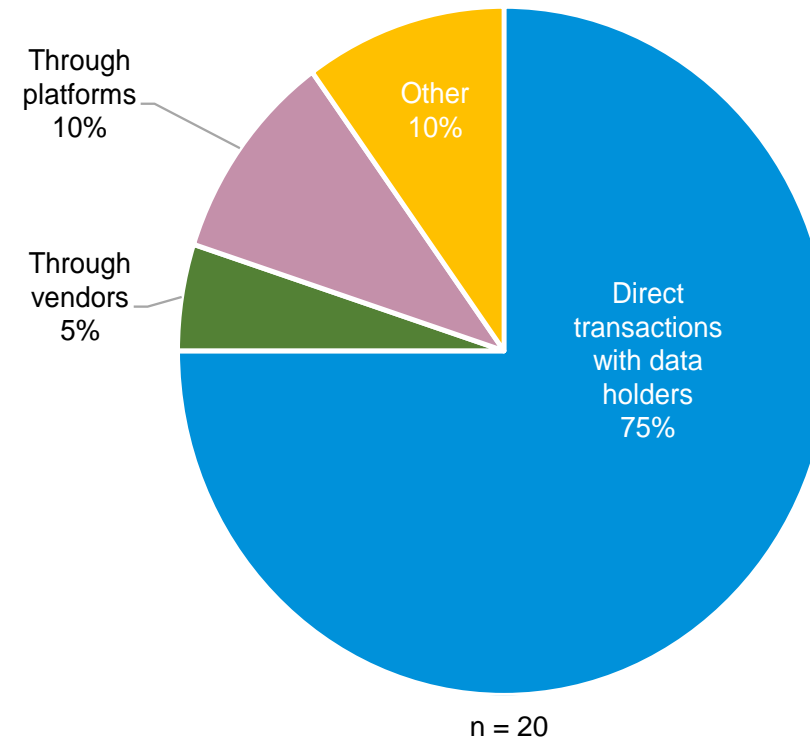
Channels of Purchasing/Selling Data

- We asked about the purchase/sales channels they use to buy/sell alternative data. The most common answer was “Direct transactions with data holders” in both buyers and data providers.
- Lots of buyers also answered “Purchase from vendors.” For both type of respondents, only limited proportions answered that they transacted via platforms.

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



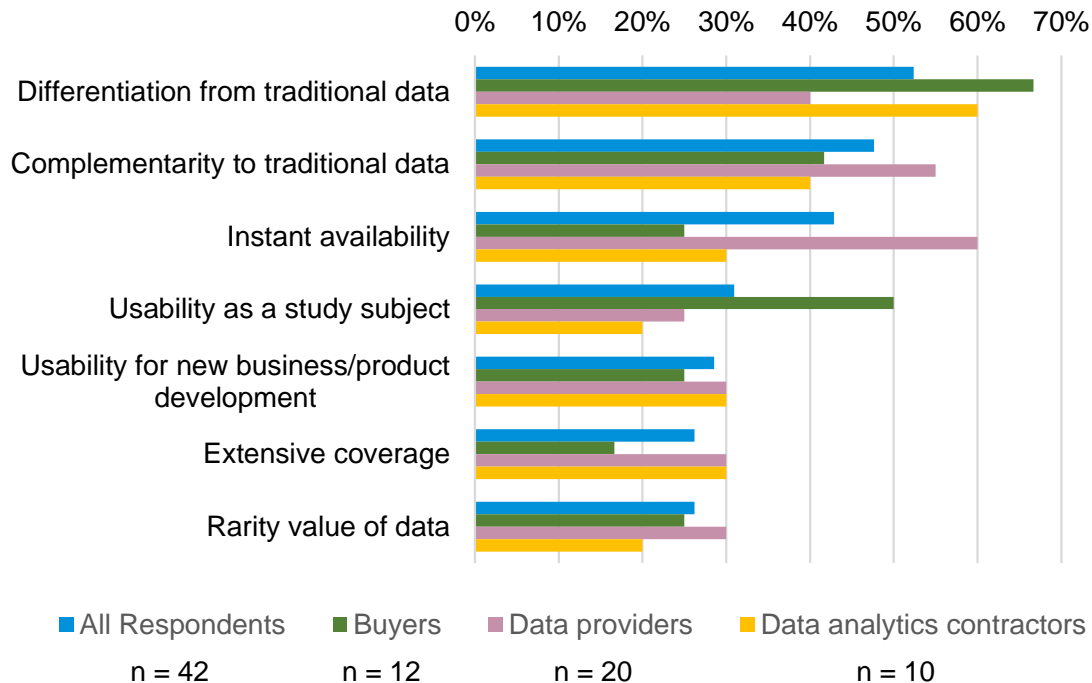
■ What sales channel do you use to sell alternative data?



Benefits of Alternative Data

- To the question about the benefits of using alternative data, 52% of all respondents selected “Differentiation from traditional data,” 48% answered “Complementarity to traditional data,” and 43% “Instant availability.”
- The result are not quite the same when responses were sorted by players. The most common answers were “Differentiation from traditional data” (67%) and “Usability as a study subject” (50%) among buyers, whereas it is “Instant availability” (60%) among data providers.

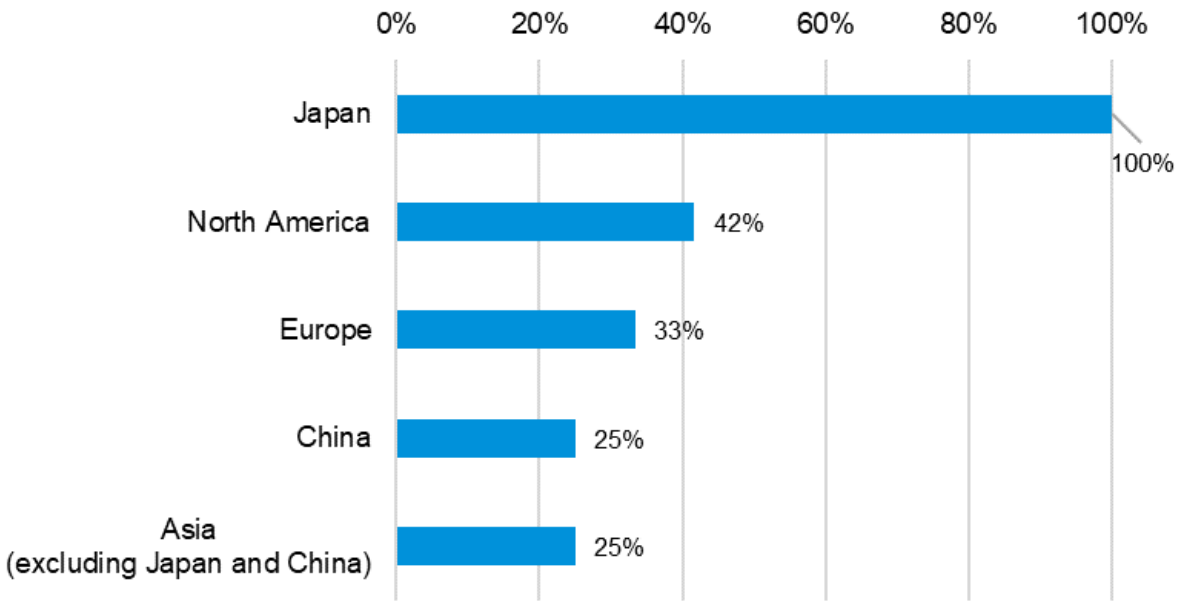
■ What is/are the benefit(s) of using alternative data? (multiple choice question, up to 3 answers)



Regions/Sectors of Alternative Data

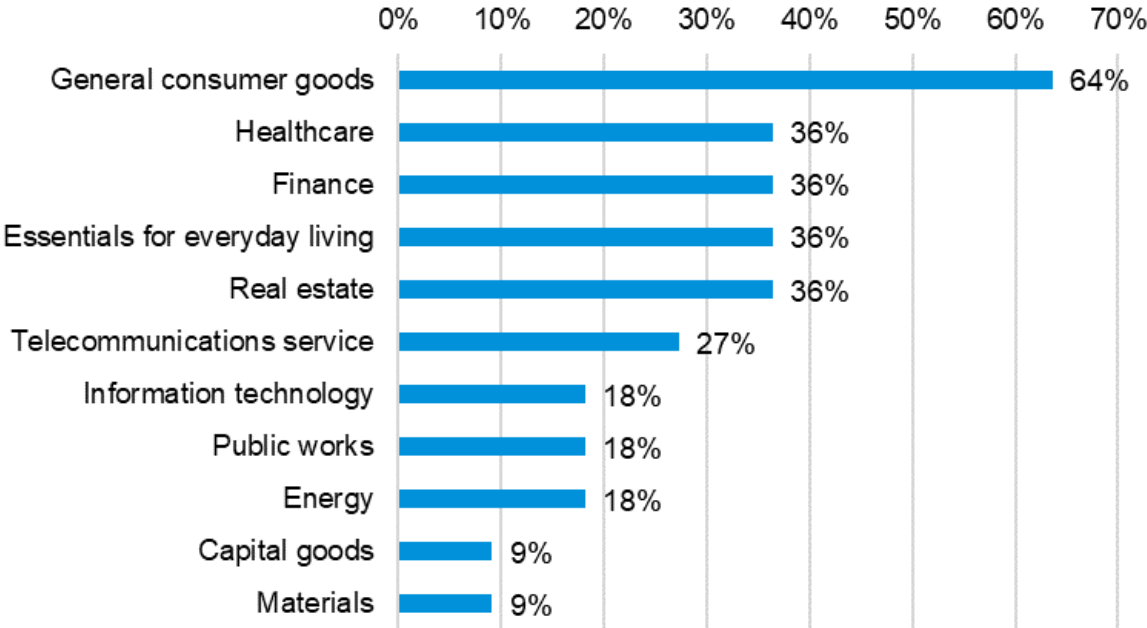
- We asked buyers the region(s) that the purchased alternative data covers. The most common answer was "Japan." For overseas, North America, Europe, China, and other Asian countries were listed, which implies that it is likely that purchased alternative data are mostly those on major economies.
- Regarding sector(s), the most common answer was "General consumer goods," followed by "Healthcare," "Finance," "Essentials for everyday living," and "Real estate."

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



n = 12

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



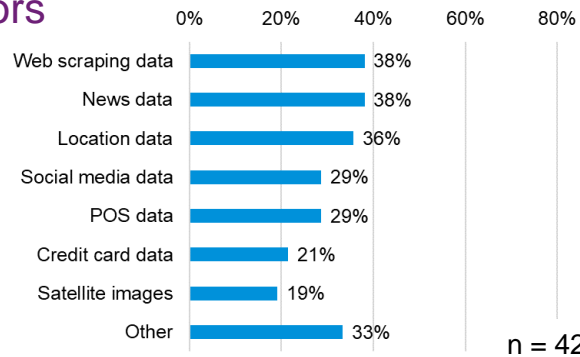
n = 12

Types of Alternative Data

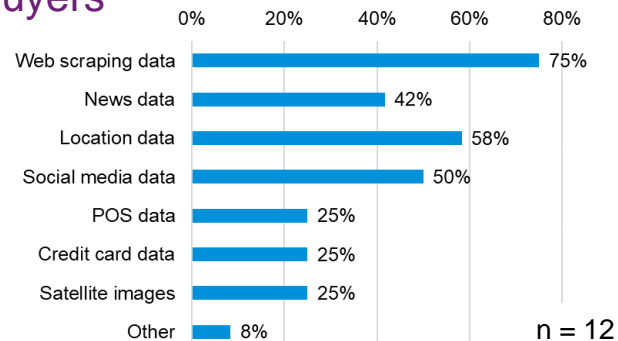
- In regard to the types of alternative data handled, most popular data for all respondents were “Web scraping data”, followed by “News data” and “Location data”.
- Buyers and data analytics contractors commonly use “Web scraping data,” “Location data,” and “Social media data,” whereas data providers were “News data” and “POS data.” In addition to these popular data, we find that various types of alternative data are used such as finance data, maritime transport data, patent-related data, receipt data, and ESG data.

Types of alternative data used (multiple choice question)

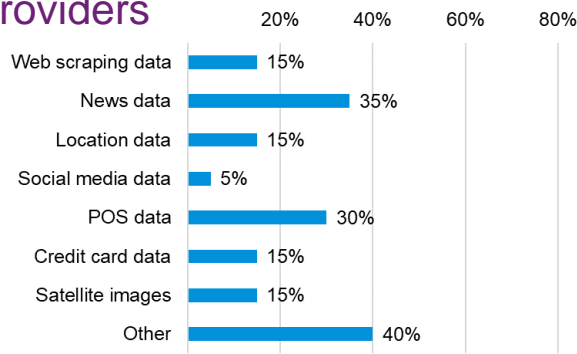
All actors



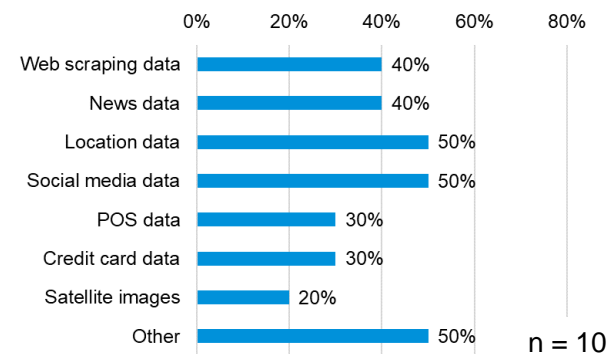
Buyers



Data providers



Data analytics contractors



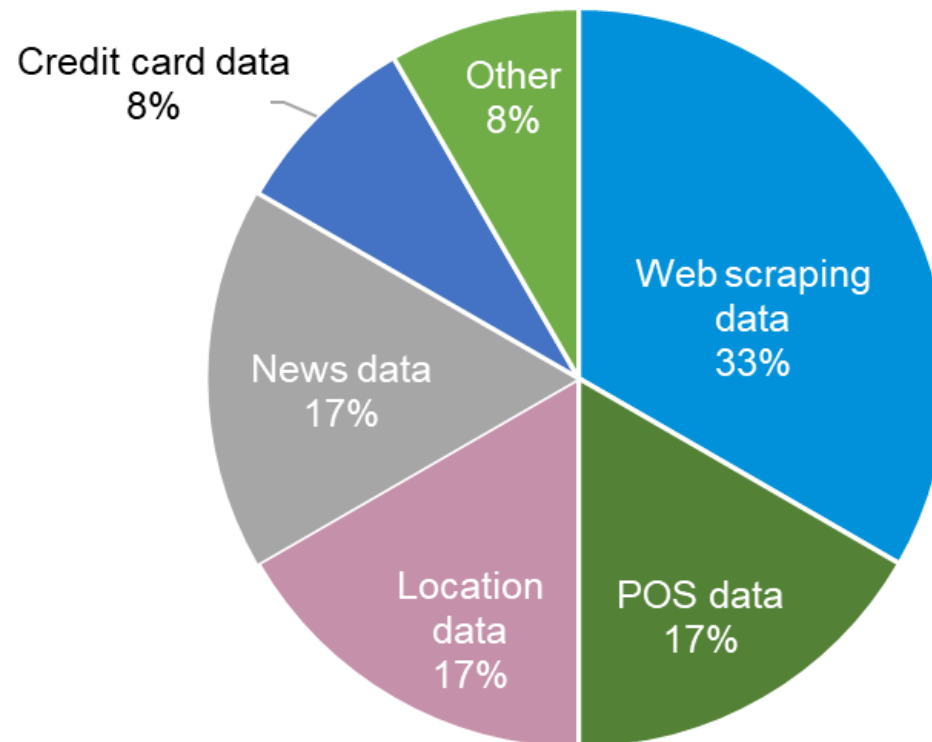
Other

- Finance data
- Maritime transport data
- ESG data
- Collected data about personnel
- Patent-related data
- Sentiment data
- Anonymous receipt data
- Investor events data

Most Frequently Used Alternative Data

- We asked buyers the type of alternative data they most frequently use. The most frequently used data is “Web scraping data” (33%). Some listed “POS data” (17%), “Location data” (17%). “News data” (17%), and “Credit card data” (8%).

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

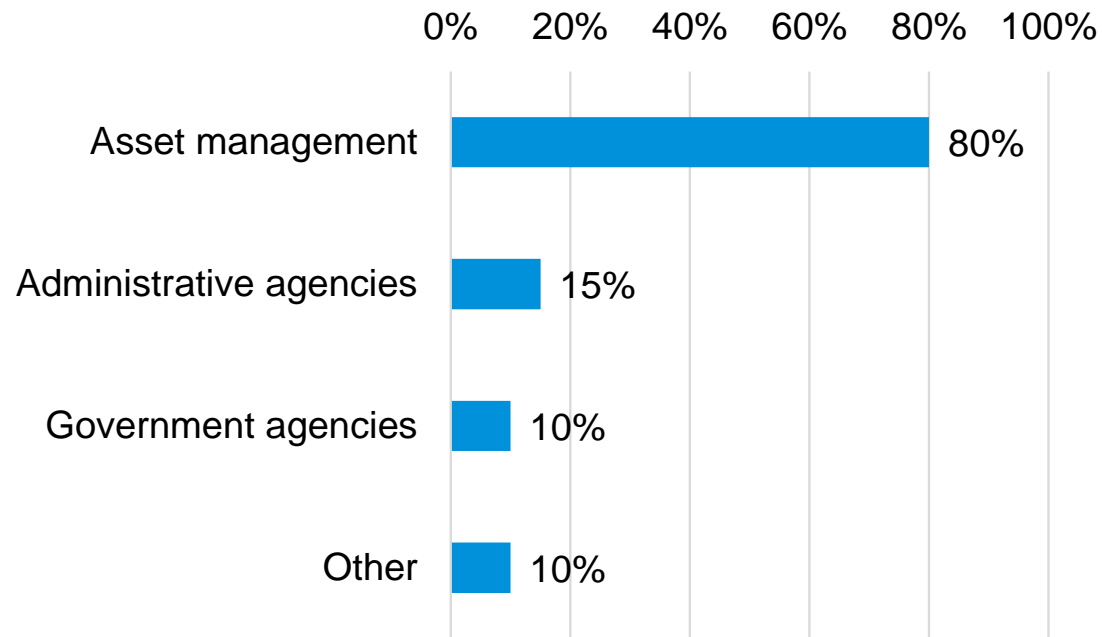


n = 12

Customers of Alternative Data Providers

- We asked data providers to whom they sell alternative data.
- The most common answer was “Asset management companies” (80%). Other answers included “Administrative agencies” (15%) and “Government agencies” (10%).

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



■ **Other**

- Universities
- Businesses
- Information vendors

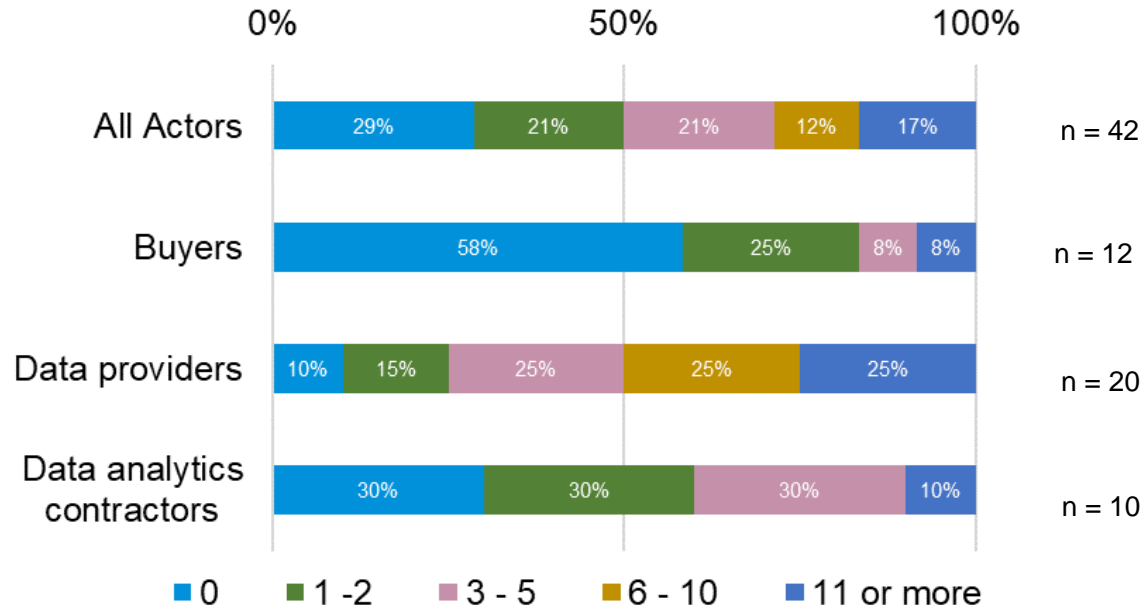
n = 20

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 an interesting polarization between the key players.
- Among buyers, 80% dedicates up to two, whereas lots of data providers and data analytics contractors dedicated more staff members.
- These result suggest that adequate systems for alternative data management are not yet developed, mainly among buyers who handle the data as users..

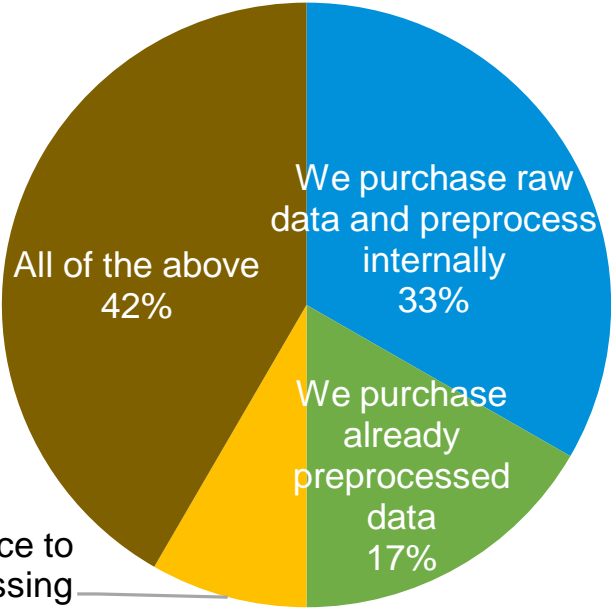
■ How many of your employees are dedicated to alternative data management?



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 or by outsourcing.
- To the question that asked the part of preprocessing, the majority answered “Remove outliers” (89%), “Remove duplicate data” (89%), and “Convert text data to numbers” (78%).

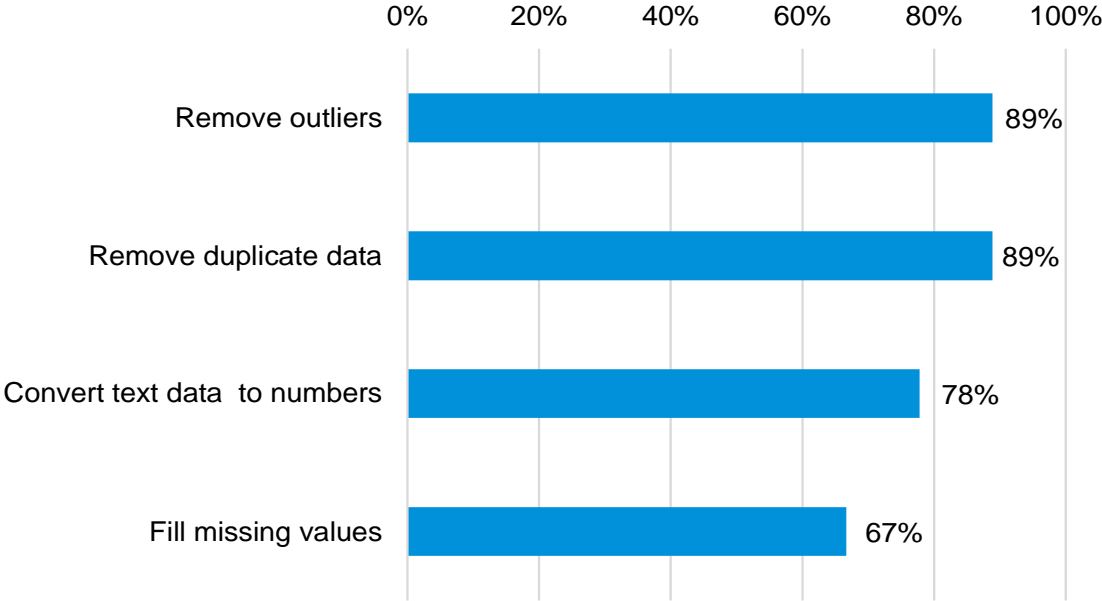
■ How do you preprocess alternative data you purchase?



We purchase raw data and outsource to an external contractor for preprocessing 8%

n = 12

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



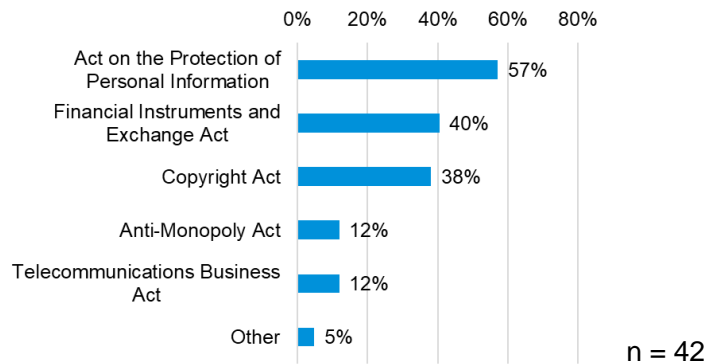
n = 9

Regulations That Need Improvement

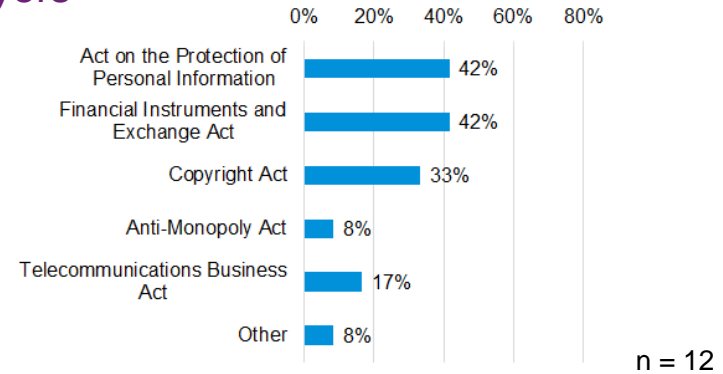
- To the question that asked about regulations that need improvement to promote the use of alternative data, the common answers from all respondents were the “Act on the Protection of Personal Information” (57%), “Financial Instruments and Exchange Act” (40%), and “Copyright Act” (38%).
- The results were all in all similar in all types of players. Showing detailed difference, data providers and data analytics contractors consider that the Act on the Protection of Personal Information are in particular needed to improve.

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

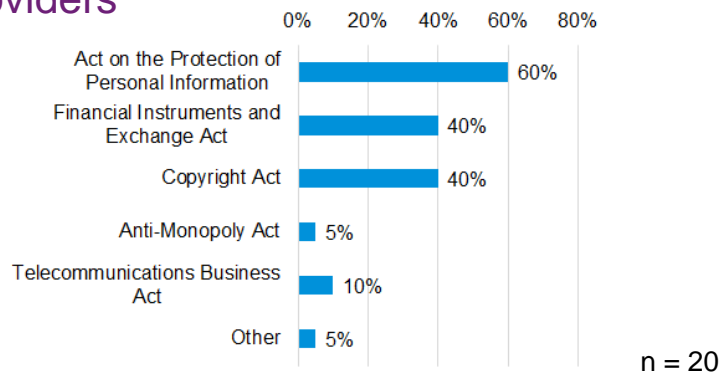
All actors



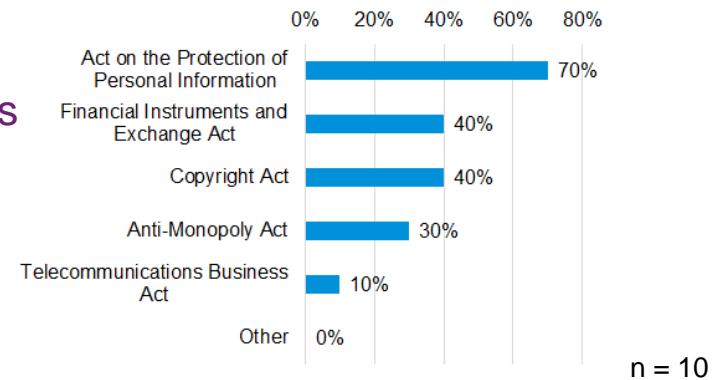
Buyers



Data providers






































Data analytics contractors



Regulations That Need Improvement (sorted by Type of Data)

- 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” is all in all believed to need improvement, regardless of the type of data handled. For the “Copyright Act”, answers varies according to the type of data used. It is commonly believed to need improvement by users of “POS data”, “Satellite data”, and “News data”, whereas the proportion were lower among users of “Web scraping data”, “Social media 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)

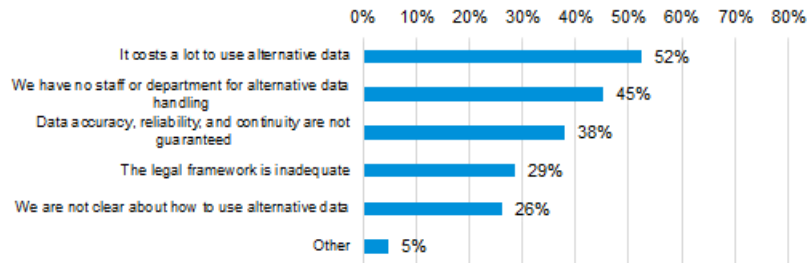
	Act on the Protection of Personal Information	Copyright Act	Financial Instruments and Exchange Act	Anti-Monopoly Act	Telecommunications Business Act	n
POS	 75%	 67%	 33%	 25%	 8%	12
Credit card	 89%	 44%	 44%	 11%	 11%	9
Location information	 73%	 33%	 33%	 13%	 27%	15
Satellite images	 88%	 50%	 38%	 25%	 0%	8
Social media	 58%	 42%	 33%	 17%	 17%	12
Web scraping	 58%	 42%	 33%	 17%	 17%	16
News	 63%	 50%	 50%	 6%	 0%	16

Issues Involved in the Use of Data

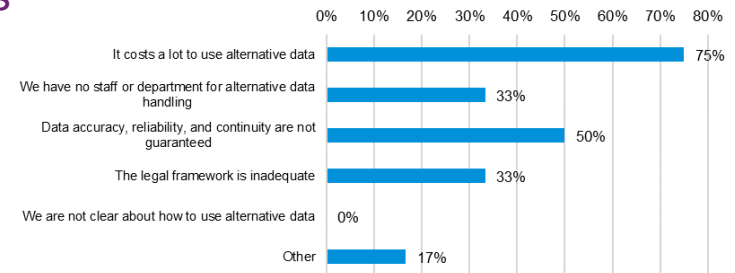
- We asked about issues involved in the use of alternative data. For all respondents, common answers were “It costs a lot to use alternative data,” “We have no staff for alternative data handling,” “Data accuracy, reliability, and continuity are not guaranteed,” and “There is no established legal framework.”
- For buyers, the most common answer was the issue of costs. For data providers, a good 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 costs, followed by issues of staff and of the inadequate legal framework.

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

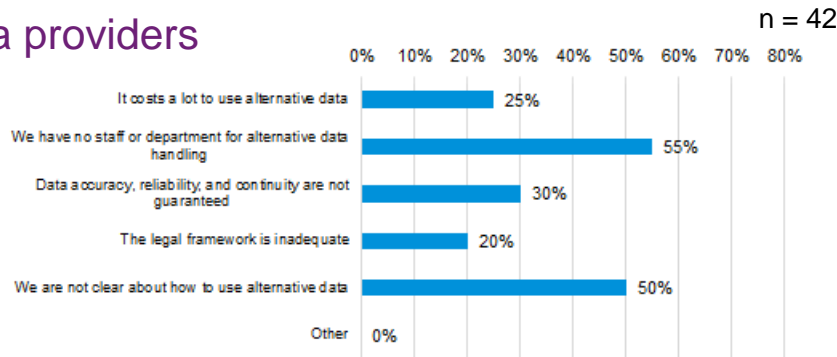
All actors



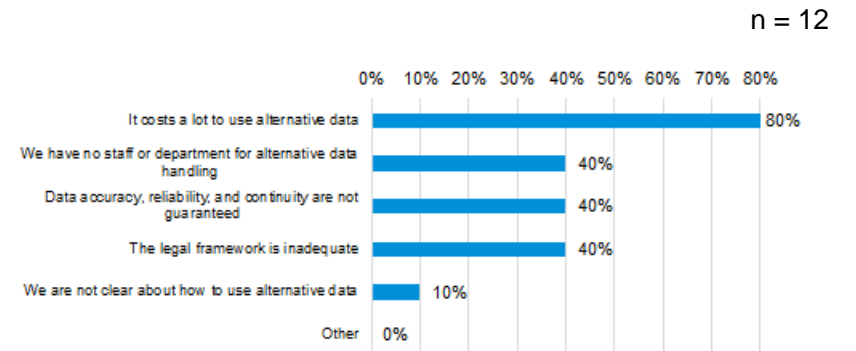
Buyers



Data providers



Data analytics contractors



Issues (Buyers)

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

We sometimes hesitate to start using alternative data because of uncertainty regarding regulations.

We are unable to take enough time to analyze alternative data.

Representativeness seems to be a major issue.

The safety of data usage is not guaranteed.

It is difficult to convincingly present the cost-effectiveness of alternative data to the management.

valuation standards for volumes and quality of alternative data.

We do not have enough data scientists.

We are unable to set up a team dedicated to alternative data handling.

We have not built up enough data to identify patterns in statistics and make seasonal adjustments.

Issues related to budgets (e.g., rising cost) are often deterrents

It is difficult to give a convincing explanation for a considerable cost before the benefits of using alternative data become visible.

Issues (Data Providers)

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

Industries and companies need to launch their own activities to advance users' understanding.

We need more employees on the staff for data handling.

Natural language processing and other relevant techniques need to be more common.

We need more employees capable of handling finance-related work.

Japan's Copyright Act puts data vendors at a disadvantage, discouraging selling data.

Ensuring data accuracy is a challenge, and legal risk is obscure.

It is difficult to convincingly present to customers exactly what value the use of alternative data offers.

Users have not acquired enough knowledge about the new type of data.

There are no established standards or third-party institution for quality assessment.

Reluctance to use a new type of data; cost-effectiveness of the data

With the sensitive nature of alternative data, the availability of the data needs to be limited.

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

Given the difficulty in clearly presenting the usefulness of alternative data, it is difficult to obtain approval for purchasing a new alternative data set.

Customers sometimes abandon their plan to use alternative data because they do not see the cost-effectiveness of the data or the data does not worth the cost.

The lead time for the use of data is long.

There are many tasks to ensure that none of the rights involved in deliverables are violated.

The accuracy of location information varies depending on the type of data.

It takes time and effort to calculate cost-effectiveness, which often prevents us from making a start, even in the stage of verification.

The data we want and laws related thereto are undeveloped.

Expensive queries are a deterrent to analyses that use multiple data sets.

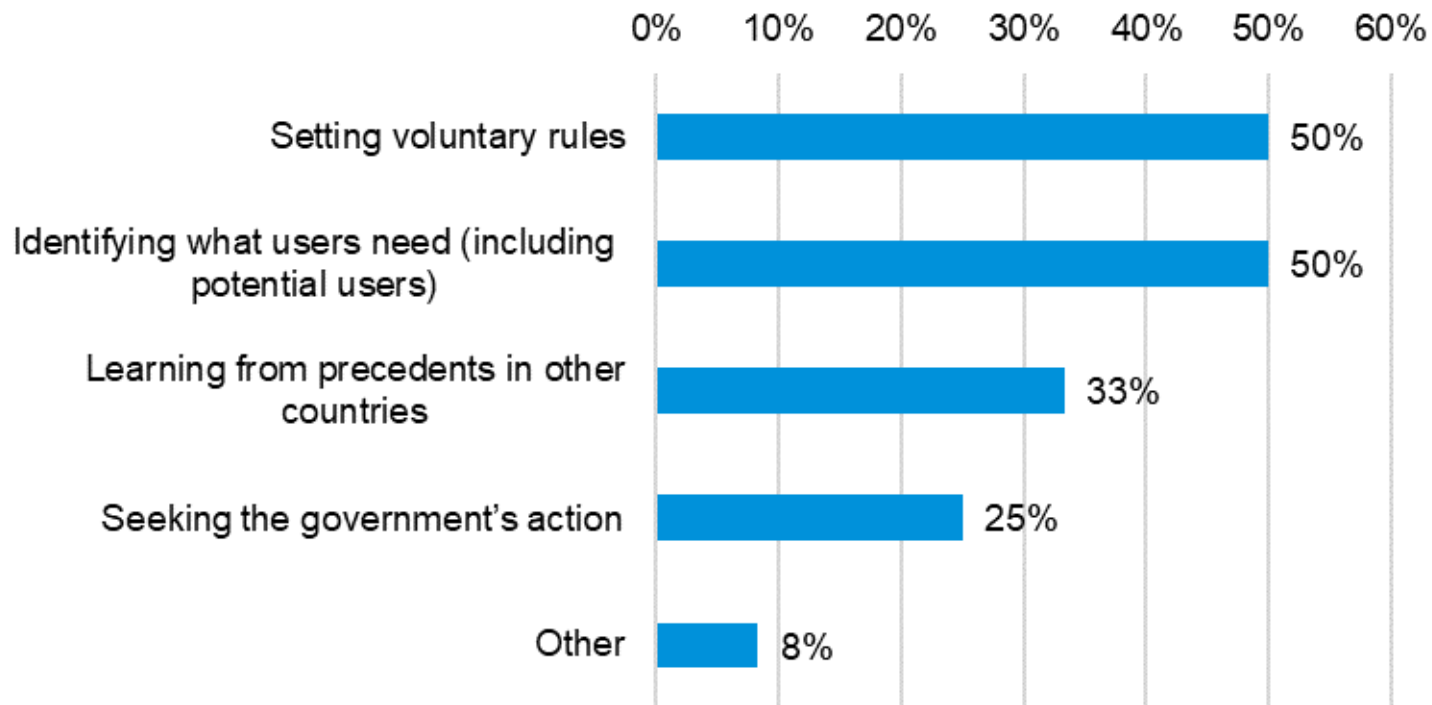
Alternative data handled at the university level is limited in Japan compared to other countries, which means alternative data experts are hardly produced.

The tough restrictions imposed by the Act on Protection of Personal Information remain a barrier.

Actions Needed to Resolve Issues

- As actions for addressing and resolving issues regarding alternative data, the respondents selected “Setting voluntary rules” (50%) and “Identifying what users need (including potential users)” (50%), and “Learning from precedents in other countries” (33%).
- On the contrary, relatively fewer respondents selected “Seeking the government’s action” (25%).

■ **What actions are needed to address and resolve the issues involved in the use of alternative data?
 (multiple choice question, up to 3 answers)**

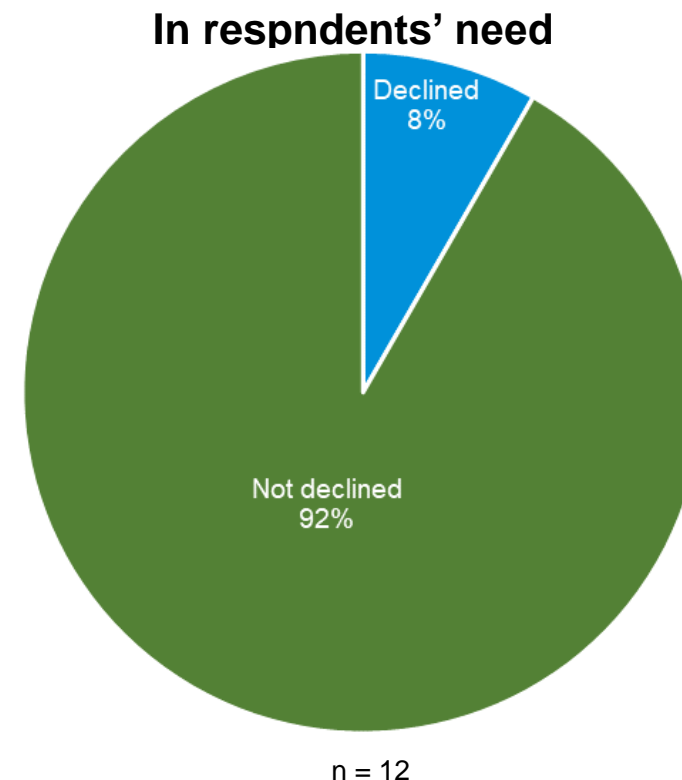
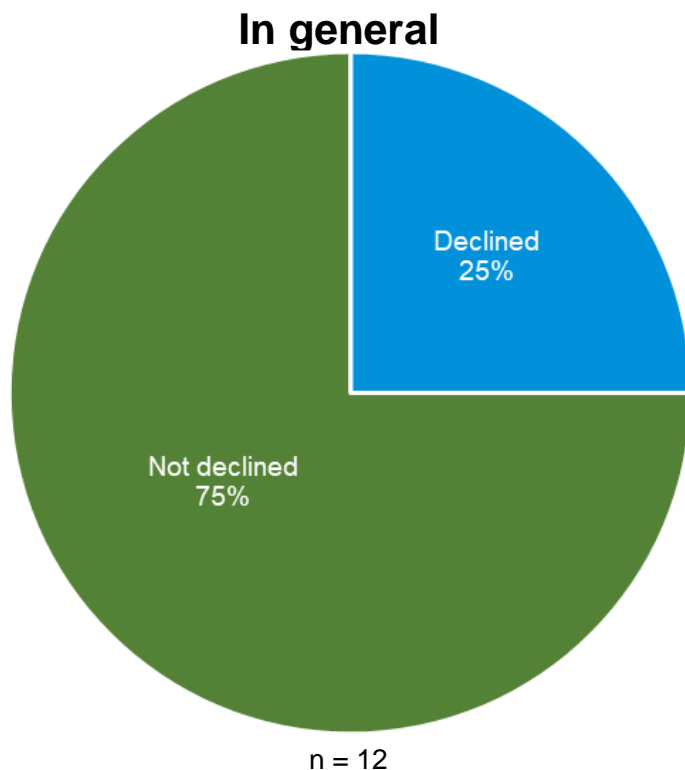


n = 12

Need for Alternative Data

- We asked buyers whether the need for alternative data had changed, both in general and in their companies, following the normalization from COVID-19.
- In general, 75% answered that the need of alternative data have not declined. In their company, the rate jumped up to 92%, indicating that the need for alternative data remains strong even after the decrease of emergent demand during the pandemic.

■ Has the need for alternative data declined in the Post-COVID economy?



Why the Need Has/Has Not Declined

■ Why the need has not declined

There is demand for POS data and web data.
There is demand for data with high granularity.
Alternative data is useful for detecting signs of what may happen in financial markets.
Alternative data is useful for quick observation of business confidence and for diversification of analytical techniques.
The use of location data and satellite data is increasingly common in financial institutions and government agencies.
We use alternative data more often than before, mostly for overseas hedge funds.
We receive more inquiries about alternative data from customers than before.
There is a growing need for foot traffic data for a business.
There is a growing need for the use of alternative data combined with map data.
Alternative data is more commonly used in the asset management business and the real estate business than before.

■ Why the need has declined

There is less need for analysis required to be instantly available.
There is less need because of budgets.
There is less need in connection with the areas of consumption we need to identify.

Outlook for Alternative Data

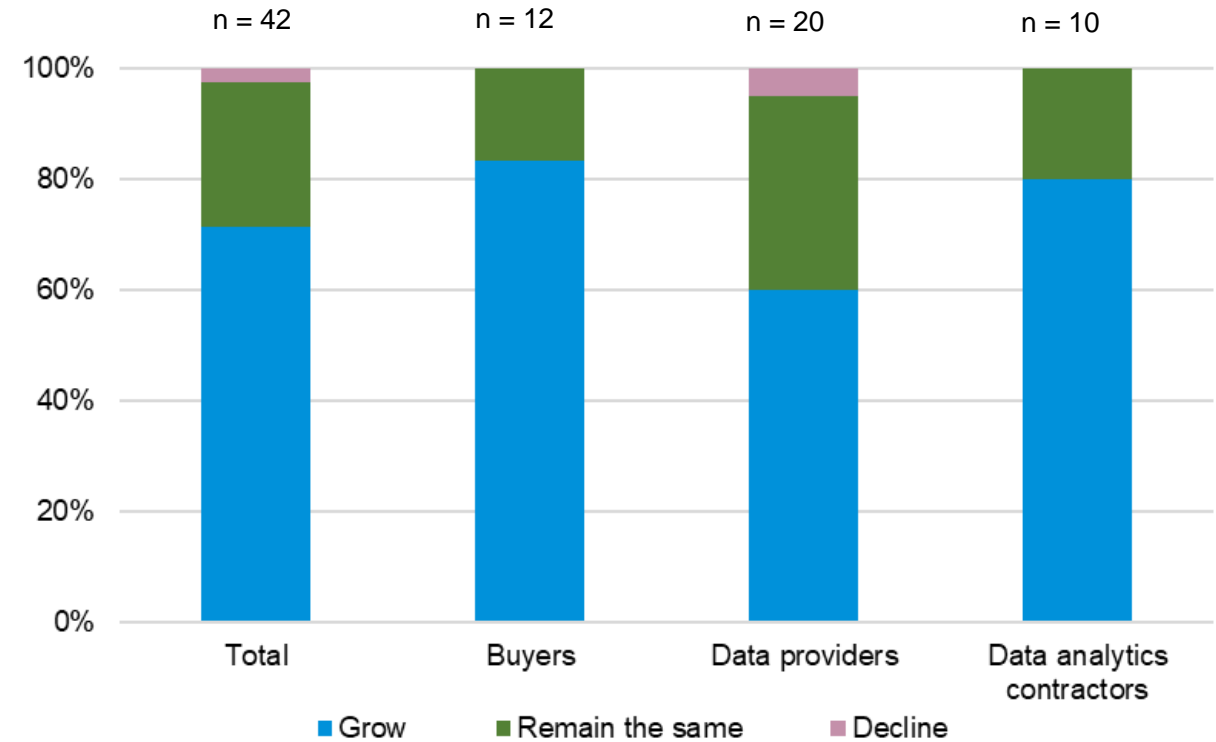
Buyers

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.”
- The results were similar regardless of key players, which indicates that respondents have high expectations for market growth of alternative data..

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

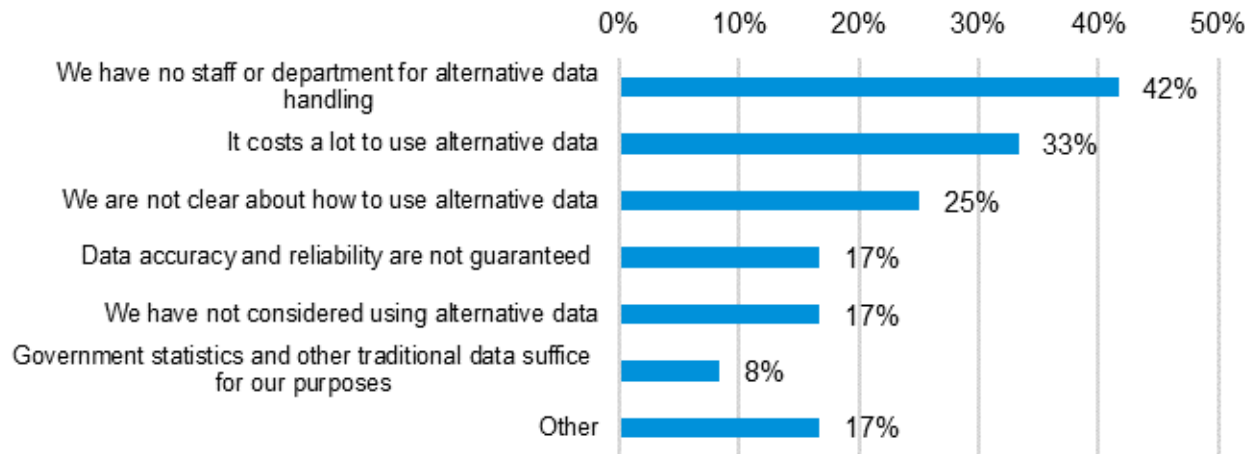


Reason for Not Using Alternative Data

Non-users

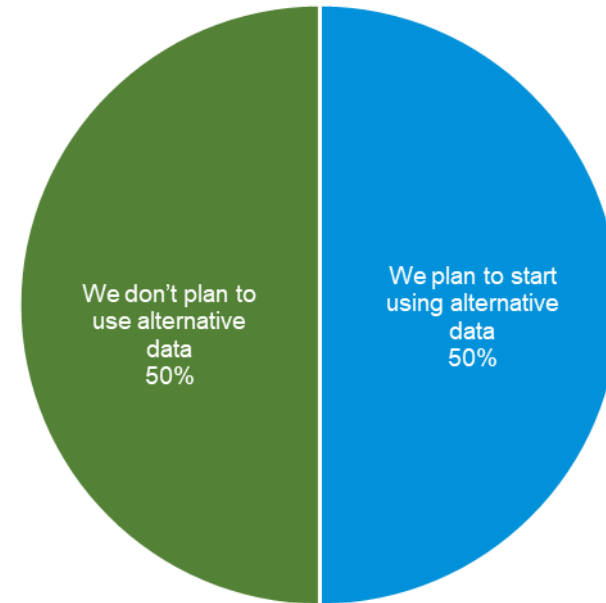
- We finally asked non-users of alternative data why they did not use it. the most selected the reason was “We have no staff or department for alternative data handling” (42%), followed by “It costs a lot to use alternative data” (33%), “We are not clear about how to use alternative data” (25%), and “Data accuracy, reliability, and continuity are not guaranteed” (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)



n = 12

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



n = 12

Conclusions

- Through this survey we confirmed that the use of alternative data is increasing.
- Even after the decrease of emergent demand during the pandemic, needs for alternative data have not declined. The results of the analyses in this report demonstrate that there is still great demand for alternative data.
- The survey found the following issues involved in promoting the of alternative data:
 - (1) It costs a lot to purchase alternative data and to develop infrastructure needed for the use.
 - (2) It is difficult to the output gained by using alternative data in advance.
 - (3) Some buyers have no staff or department to handle alternative data.
 - (4) Data accuracy, reliability, and continuity are not guaranteed.
 - (5) The legal framework is inadequate.

We hope to continue this survey regularly to study trends in the use of alternative data, if members of the JADAA would agree to participate again. 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.