International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

To What Extent Does Digital Data Analysis Benefit Small and Medium-Sized Businesses (SMBs) in Japan?

Mei Hanaoka

Student researcher, Grade 12, Lancers International School

Abstract
Digital data analysis has become a key factor in adopting digital transformation, providing valuable insights based on logic and empirical evidence to navigate decision making processes in today's rapidly changing industries. However, its implementations in Japanese small and medium-sized businesses (SMBs) which play an indispensable role in the Japanese economy remains limited, primarily due to its unconvincing benefits compared to current traditional methods. Therefore this paper aims to evaluate the extent to which digital data analysis benefits SMBs in Japan, taking into consideration aspects such as online and in-store purchasing behavioral patterns and employee communication.

Keywords: Data Science, Small and Medium-Sized Businesses (SMBs), Digital Transformation

1. Introduction

1.1 Background Information

Referring to the collection and analysis of big data in order to enhance business efficiency and productivity, data analysis has now become a necessity for businesses of all fields and sizes (HCLTech, n.d.), extending its reach to traditional small and medium-sized businesses (SMBs) in Japan. Although some companies prefer to hire full-time specialists to develop and manage data analysis systems to provide business insights and strategies, in recent years data analysis has become much more accessible for SMBs. Options such as part-time consultants who regularly guide the company based on data-driven decisions and digital analysis which utilizes software to make data analysis accessible to business owners with no prior knowledge (business.com, n.d.) enables SMBs to adopt data-driven initiatives for a smaller budget, increasing their competitiveness in today's rapidly changing industries.

SMBs are defined as businesses with less than 300 employees (Statista, n.d.), enabling them to excel in flexibility and personalised customer service while also increasing the need for high efficiency and productivity to fully operate within their limited resources. They play a crucial role in sustaining a country's economy, accounting for around 90% of all businesses and creating two thirds of the world's employment (WhatIs, 2023). Moreover, this trend is especially prominent in Japan with over 3.36 million SMBs, accounting for 99.7% of all businesses and playing an indispensable role in terms of employment creation (World Economic Forum, 2025).

However, SMBs lack the resources required to adopt large-scale IT solutions utilized by larger enterprises, with high initial costs and compatibility issues with both employees and existing systems delaying its widespread adoption. Based on statistics, only 55.6% of SMBs conduct data analysis compared to the

90.9% for large companies, with 80% answering that they only view the data to understand trends rather than conducting further quantitative evaluations (Data no Jikan, n.d.).

Furthermore, this delay in the SMBs' digital transformation is also due to many business owners finding these excessive data-based analysis unnecessary, considering their business size which enables them to intuitively grasp company sales trends through hands-on experience and consumer interactions (Implement, 2021). As a result, Japan is said to have the lowest digital maturity score compared to neighboring nations such as Singapore and Hong Kong (Sia Partners, 2019), with a survey conducted for 1,500 executives of SMEs in China, Indonesia and Japan showing how 72% of the SMBs with no digital sales do so due to its unconvincing business benefits (Consultancy.asia, 2019). Nevertheless, with the expansion of sales to online platforms and rising emphasis on employee communication, these traditional methods subject to biases may lose its prior effectiveness in today's data-driven business world. As mentioned in the book *'Rewired: A McKinsey Guide to Outcompeting in the Age of Digital and AI*', digital transformation has now become indispensable in establishing a competitive differentiation through lower costs and enhanced customer experience (McKinsey & Company, 2024).

Therefore this paper aims to evaluate ways in which digital analysis, which utilizes software to make data science more accessible to non-experts, can coexist with existing decision making processes adopted by most SMBs in Japan, acting as a tool for guidance rather than the full employment seen in larger enterprises. The discussion will be centered around the research question "*To what extent does digital data analysis benefit small and medium-sized businesses (SMBs) in Japan*?", taking into consideration realistic benefits of data analysis for SMBs through an observation of a local traditional SMB *K* located in Japan. The next section will be analysing the data derived by company *K*, and identifying ways in which digital analysis may provide SMB owners with previously unknown business insights and strategies. The last section will evaluate the extent in which these insights may benefit the SMBs, as well as mentioning the research's limitations and possible future extensions.

1.2 Research Method

To further understand the realistic benefits for SMB owners to implement digital data analysis, I interned at a Japanese traditional SMB *K* for 2 weeks, analyzing daily sales reports and having regular online meetings with the CEO regarding my findings on customer purchasing behavioural patterns and supplier performance. I utilized a free online data analysing software called JupyterLab which is an interactive development environment (IDE) using a command line interface (CLI), essentially meaning it enables code to be tested and visualized in 'notebooks' through the programming language Python. Through this software I ran commands related to data cleaning, arithmetic calculations and data visualization, deriving values and diagrams which I later included in my daily and weekly reports. My main aim was to search ways in which digital data analysis could provide insights previously unattainable by a single owner, as well as search for other benefits in quantifying intuitive understandings.

```python
# Import necessary libraries to interpret the pdf
import pdfplumber
import pandas as pd

all_rows = []

# Open the pdf '0616.pdf' stored in the same file, and extract its data table
with pdfplumber.open("0616.pdf") as pdf:
    for page in pdf.pages:
        table = page.extract_table()
        if table:
            all_rows.extend(table[1:])  # Skip headers on each page

# Use the first page's header as the dataframe's header
header = pdf.pages[0].extract_table()[0]
df = pd.DataFrame(all_rows, columns=header)

# Rename necessary columns
df.columns = ['serial code', 'name/price/size', 'sales1/inStock/sales2','order', 'quantity/revenue', 'supplie
```

Figure 1. Python Code to Extract Tables

```python
# Import necessary libraries for data visualisation
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Plot a boxplot for the profit gained from each customer with range -100 to 4500 yen
sns.boxplot(x=grouped['profit'])
plt.title('Profit')
plt.xlim(-100, 4500)
plt.show()

# Plot a boxplot for the revenue gained from each customer with range -100 to 7000 yen
sns.boxplot(x=grouped['revenue'])
plt.title('Revenue')
plt.xlim(-100, 7000)
plt.show()
```

Figure 2. Python Code to Plot Box Plots

```python
# Group inStore data in terms of individual products sold
grouped_inStore = (
    inStore_df
    .groupby(['serial code'])
    .agg(
        product_name = ('product name', 'first'),
        total_quantity=('quantity', 'sum'),
        row_count=('quantity', 'size'),
        total_profit=('profit', 'sum'),
        price=('price', 'first'),
        total_revenue=('revenue', 'sum')
    )
    .reset_index()
)
```

Figure 3. Python Code to Separate In-store and Online Purchases

The daily sale reports sent as a pdf were read and converted into a dataframe through the pdfplumber and pandas library, enabling the analysis of the data stored through Python commands. Figure 1 shows an example code for data cleaning where the data is retrieved from the pdf and the table columns are renamed; Figure 2 plots a box plot for the profits and revenue by each customer using the seaborn and matplotlib.pyplot library to visualise the distribution of data; and Figure 3 separates in-store and online

purchases and creates a new table in terms of products sold which includes product name, total quantity, total revenue, and more.

Moreover, secondary sources such as literary searches of both Japanese and English websites and articles were also utilized to support the proposed arguments, gaining information regarding Japanese SMB's delayed adoption of data analysis, common issues faced and possible benefits of data visualisation.

2. Findings and Analysis

2.1 Online Purchases

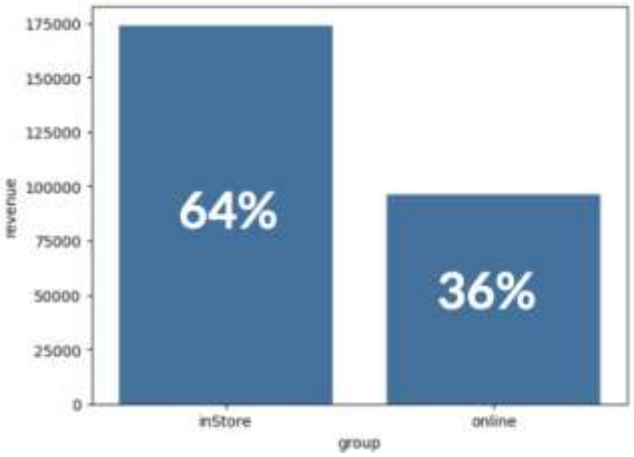

Figure 4. Daily Revenue for In-store Purchases and Online Purchases

Figure 5. Bar Graph Comparing In-store and Online Average Purchasing Quantity, Total Purchase Price and Product Price

```python
import pandas as pd
df = df.copy()

# Ensure 'receipt number' is string
df['receipt number'] = df['receipt number'].astype(str)

# Create new 'group' column to differentiate in-store and online purchases
df['group'] = 'Other'

# Store 'inStore' to column 'group' for purchases where the receipt number starts with 1
df.loc[df['receipt number'].str.startswith('1', na=False), 'group'] = 'inStore'

# Store 'online' to column 'group' for purchases where the receipt number starts with 3 or 4
df.loc[df['receipt number'].str.startswith(('3', '4'), na=False), 'group'] = 'online'
```

Figure 6. Python Code to Differentiate In-store and Online Purchases

One of the ways in which digital data analysis provided a new insight to the CEO was through the comparison of in-store and online purchases. By separating the two genres into different tables and plotting bar graphs to compare their total daily revenue proportions and the characteristics of each consumer purchase, a clearer contrast between the two categories was enabled.

As seen in Figure 4 online purchases make up a smaller portion of total revenue at 36% compared to the 64% by in-store purchases. However, Figure 5 shows how the average quantity, total price and average price per purchase is much greater for online compared to in-store visitors, indicating online shoppers' larger individual purchase sizes. Furthermore online customers had close to no overlaps in consuming patterns in contrast to in-store purchasers, most likely as they are less affected by offline marketing strategies such as product positioning and in-store events, and rather shop with a clear purpose to purchase a specific product.

Although the owner was intuitively aware of online customers generally belonging to the wealthy class such as professional artists and experts through consumer profiles, they also lacked empirical evidence to clearly confirm these trends in sales. This was due to online and in-store purchases being indicated only by the first digit of the receipt number, therefore being quite difficult to differentiate through pure observation of the data. However, with digital data analysis the two categories could easily be separated with the code shown in Figure 6, enabling a more comprehensive analysis and revealing previously unknown patterns.

In recent years due to factors such as increased accessibility and the COVID-19 pandemic, the adoption of online platforms by SMBs in Japan are steadily increasing (SME Japan, 2023), with e-commerce expected to grow at double the rate of in-store purchases between 2022 and 2027 (Agriculture and Agri-Food Canada, 2025). Although traditional methods often prove to be quite accurate due to the employees' rich first-hand experience, this strategy may lack competitiveness for online sales which are harder to intuitively grasp with minimal in-person interactions. Therefore digital data analysis which can logically differentiate the two categories and analyze them separately enables more accurate comparisons, providing key insights in expanding a business's e-commerce.

2.2 Data Trends Over Time and Biases

```python
def outing_comfort_score(temp, humidity, wind, precipitation_mm):
    score = 0

    # Temperature (ideal: 23°C)
    score += max(0, (1 - abs(temp - 23) / 15) * 3)

    # Humidity (ideal: 50%)
    score += max(0, (1 - abs(humidity - 50) / 50) * 2)

    # Wind (ideal: ~3 m/s)
    score += max(0, (1 - abs(wind - 2) / 6) * 1)

    # Precipitation amount (non-linear scoring)
    score += precipitation_score(precipitation_mm)

    return round(score, 2)
```

Figure 7. Calculation Algorithm for Outing Index

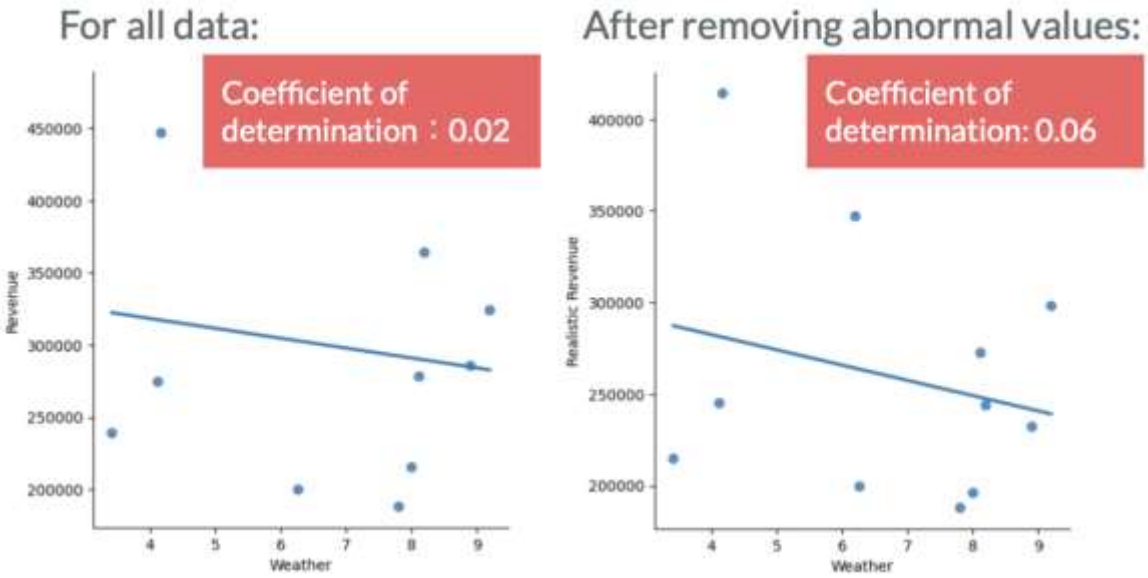

Figure 8. Scatter Diagram of Outing Index and Daily Revenue

Another aspect in which digital analysis provided new insights for the owner was the analysis of correlation between total daily revenue and weather. The owner speculated how the number of customers and therefore the day's total revenue decreased during bad weather and increased on sunny days more fit for outings. To confirm this hypothesis the daily sales reports and official weather reports provided by the Japan Meteorological Agency over the two weeks were used for a quantitative evaluation.

In order to numerically identify their relationship an 'outing index' was calculated using the algorithm shown in Figure 7, which considered the day's temperature, humidity, wind and precipitation to calculate a representative score of how fit the weather was to go out. A higher score indicated a more comfortable weather, and these values were plotted on a scatterplot along with the day's total revenue as seen in Figure 8 to observe their correlation. In either case of including or excluding bulk purchases, the coefficient of determination also known as R squared was notably low at 0.02 and 0.06, indicating an extremely weak relationship between weather and daily revenue and thus disproving the owner's original speculation.

As changes in revenue and purchasing patterns over time are complex to grasp simply through first-hand observations and suitability for outings are a subjective topic differing for each individual, numerically expressing these perceptive concepts allowed for a more objective analysis. The quantitative data analysis was free of bias held by the owner's personal experience of being discouraged to go out during bad weather, providing a more accurate understanding of the correlation between weather and daily revenue.

2.3 Employee Communication

Furthermore, digital data analysis simplifies complex raw data to processed information, enabling smooth communication of sophisticated company sale trends initially limited to experienced and high-post individuals. Traditional methods of simply viewing the raw data for analysis requires experience and a comprehensive understanding of the business's background, thus making it exclusive for senior members. In contrast, data visualisation projects the data to graphs and diagrams which are easily interpreted by all employees, therefore democratizing the decision making process and promoting collaboration (Francis Xavier Engineering College, 2024).

In reality, high career-turnover rates in SMBs have become a prominent issue in recent years, with SMBs having a 3.7% higher turn-over rate in comparison to larger companies. This labour shortage has caused

350 business bankruptcies in Japan in 2024, with the lack of employee retention and excess need for recruitment and retraining hindering business productivity (World Economic Forum, 2025). In addition, research reveals lack of employee engagement is one of the leading causes of career turnovers for small businesses, caused by the employees' disconnection from the company's business goals and visions (Metrobi, n.d.). In reality, engaged employees are proved to be more productive and loyal thus with less likelihood to leave. However, in traditional SMBs with crucial discussions exclusive to certain top posts, employees are left with monotonous and unchallenging work with limited opportunities to enhance and develop their skills, thus leading to demotivation and lack of employee engagement.

This is why digital data analysis plays a key role in enhancing employee satisfaction, facilitating effective communication of complex business trends through data visualization. By sharing business data through an organised report it ensures employees are all on the same page, providing a clear sense of company direction while also drawing new insights which help build consensus around key decisions. As a result this unity enhances productivity, increasing employee engagement and thus preventing the creation of autocratic businesses.

2.4 Generational Transfer

The scalability of digital analysis in comparison to traditional methods is also a key benefit of its implementation, facilitating generational transfer of businesses by making data analysis more accessible. In fact, one of the leading causes for young generations in Japan becoming unforthcoming to take over family business lies in the lack of clarity both in the company's future plans and operational responsibilities (Payne Hicks Beach, 2024). With many SMBs relying on intuition-based decision making processes dependent on first-hand experience and background knowledge, new owners lacking these characteristics often face difficulty in taking over (Sample HubSpot User, n.d.). On the other hand, digital data analysis is scalable and can be continuously utilized by future generations with minimal changes, making it more approachable for new-comers compared to traditional qualitative methods. Furthermore, making key decisions through a unanimous agreement increases transparency in the company's future vision, enabling SMBs to adapt to the increasing competition accelerated through globalisation.

Currently in Japan the lack of successors for SMBs have become a significant issue, with a 2024 survey revealing 52.1% of Japanese companies don't have a successor to take over the business (World Economic Forum, 2025). As a result even profitable enterprises have been forced to shut down, leading to approximately 6.5 million job losses by 2030 and creating a devastating impact on Japan's economy (LinkedIn, 2025). Despite the government and private organisations' efforts to connect traditional SMBs and entrepreneurs, its influence remains limited with many young family members unwilling to take over their family's SMBs. However, the use of technology along with traditional methods can increase transparency of company goals and decision making processes, thus making data analysis accessible for young generations with less experience and facilitating generational transfer.

3. Limitations and Future Extensions

Although a thorough literacy search was conducted through Japanese and English websites, my primary research is limited to the data and insights gained from a single small-scale family business, possibly hindering its comprehensive representativeness of Japan's overall SMB population. Moreover, the full scope of the company *K*'s big data was not made available to me due to data security concerns, thus limiting me to analyze digital data's benefits only through four key aspects. Therefore in terms of future extensions, a further disclosure of data from past years may enable a more comprehensive analysis,

especially regarding the changes over time in in-store and online purchase behaviour as well as the transformation of overall sale trends before and after the company's implementation of social media marketing. Although due to my prior knowledge in programming I opted for the CLI interface IDE JupyterLab for this investigation, other more user-friendly softwares such as Tableau and Power BI may also be used for data visualisation and analysis by business owners with no experience in coding (Coursera, 2023).

4. Conclusion

As British mathematician Clive Humby mentioned in 2006, 'Data is the new oil' and thus has the ability to power business innovations if processed and utilized correctly, just like how oil can power engines when refined and processed as gasoline (Humby, n.d.). The adoption of digital transformation has become a prerequisite in recent years regardless of the business's scope or industry, with its first step being digital data analysis deeply engrained in the fundamental decision making process. Considering SMBs strengths in flexibility and rich first-hand experience, this fusion of traditional methods and upcoming technologies can enable them to swiftly adapt and thrive in today's rapidly-changing industries, acting as a tool for aid rather than a threat of replacement. This investigation revealed how digital data analysis can provide bias-free insights previously unattainable by simple observations of the data, as well as empower employees by democratising the decision making process which as a result facilitates generational transfer. Therefore despite possible setbacks both from financial and cultural aspects, digital analysis is now becoming a key factor in business survival, navigating decision making processes through logical and data-based insights.

References

1. 15 data analysis tools and when to use them. (2023, August 22). Coursera. https://www.coursera.org/articles/data-analysis-tools#
2. E-Commerce unleashed: Fueling growth in Japan's SMEs. (2023, April 16). SME Japan | Business in Japan. https://www.smejapan.com/business-news/ecommerce-unleashed-fueling-growth-in-japans-smes/
3. Statista. (n.d.). Statista - The Statistics Portal for Market Data, Market Research and Market Studies. https://www.statista.com/topics/13372/smes-in-japan/
4. How small businesses can take advantage of big data. (n.d.). business.com. https://www.business.com/articles/data-analysis-for-small-business/
5. Humby, Clive. (n.d.). The University of Sheffield. https://sheffield.ac.uk/cs/people/academic-visitors/clive-humby#
6. Implement, T. (2021, November 18). Why some businesses are slow to adopt new digital technologies? Medium. https://medium.com/@YourCRMPartner/why-some-businesses-are-slow-to-adopt-new-digital-technologies-8505155fc28b
7. Japan faces a looming crisis: The quiet collapse of its small businesses. (2025, June 11). LinkedIn. https://www.linkedin.com/pulse/japan-faces-looming-crisis-quiet-collapse-its-small-businesses-sok-1xotc
8. "Japan's Smes Are Receiving a Recruitment and Retention Boost." World Economic Forum, www.weforum.org/stories/2025/05/japan-small-businesses-economic-resilience/. Accessed 27 July 2025.

9. "Japan's Succession Problem: How the Country Is Safeguarding Heritage through Business." World Economic Forum, www.weforum.org/stories/2025/04/japan-business-succession-traditional/. Accessed 27 July 2025.
10. Metrobi.com | 520: Web server is returning an unknown error. (n.d.). Metrobi Local Delivery Courier Service for Businesses. https://metrobi.com/blog/common-causes-of-employee-turnover-at-small-businesses/
11. The role of data visualization in business intelligence: Driving informed decision-making | | Francis Xavier engineering college, Tirunelveli. (2024, February 24). The Role of Data Visualization in Business Intelligence: Driving Informed Decision-Making | | Francis Xavier Engineering College, Tirunelveli. https://www.francisxavier.ac.in/blog/the-role-of-data-visualization-in-business-intelligence-driving-informed-decision-making
12. "Tips for Achieving a Successful Transfer of a Family Business to the next Generation." Payne Hicks Beach, 22 Oct. 2024, www.phb.co.uk/article/tips-for-achieving-a-successful-transfer-of-a-family-business-to-the-next-generation/.
13. Sample HubSpot User. (n.d.). 後継者不足とは？6つの原因を解説！. ブログ | 中小企業の経営サポートなら古田土経営・古田土会計. https://blog.kodato.com/lack-of-successor
14. Sector trend analysis – e-Commerce market trends in Japan - agriculture.canada.ca. (2025, July 2). Language selection - Agriculture and Agri-Food Canada / Sélection de la langue - Agriculture et Agroalimentaire Canada. https://agriculture.canada.ca/en/international-trade/market-intelligence/reports-and-guides/sector-trend-analysis-e-commerce-market-trends-japan#
15. Sia partners ranks Singapore brands as the best digital performers. (2019, August 13). Consultancy.asia | Asia consulting industry platform. https://www.consultancy.asia/news/2417/sia-partners-ranks-singapore-brands-as-the-best-digital-performers
16. SMEs in Indonesia, China and Japan remain reluctant on digital adoption. (2019, August 27). Consultancy.asia | Asia consulting industry platform. https://www.consultancy.asia/news/2456/smes-in-indonesia-china-and-japan-remain-reluctant-on-digital-adoption
17. What is an SMB (Small and medium sized business or small and midsize business) | Definition from tec. (2023, 20). WhatIs. https://www.techtarget.com/whatis/definition/SMB-small-and-medium-sized-business-or-small-and-midsized-business
18. "What Is Digital Analytics?" HCLTech, www.hcltech.com/knowledge-library/digital-analytics-overview. Accessed 27 July 2025.
19. "What Is Digital Transformation?" McKinsey & Company, McKinsey & Company, 7 Aug. 2024, www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-digital-transformation#.
20. 日本企業におけるデータ活用の実態！ 海外と比較したときの特徴とは？. (n.d.). データで越境者に寄り添うメディア データのじかん. https://data.wingarc.com/data-utilization-in-japan-49163