Proceedings of the 4th International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/2025.19806

A Case Study on Apple's Use of Big Data to Enhance Supply Chain Transparency

Shiyue Gao[1,a,*]

[1]*Beijing Institute of Technology, College of Mathematics Physics and Civil Engineering, Zhuhai, 519088, China*
a. 13614630234@163.com
**corresponding author*

Abstract: With the continuous development of information technology, big data has become an important tool to enhance supply chain transparency. As a globally renowned technology enterprise, Apple's supply chain management has an important influence in the world. Apple has a supply chain covering multiple suppliers, and there is an increasing demand for supply chain transparency. This paper uses principal component analysis to perform dimensionality reduction analysis on the operating revenue and growth rate data of some suppliers in Apple's supply chain. The findings show that principal component analysis can effectively extract the core characteristics of suppliers and classify them into different levels. Some suppliers (e.g. Wintech) have outstanding performance, while others (e.g. Shengli Precision) show that have certain risks. This study aims to provide a reference for other enterprises, especially multinational corporations, and to help them enhance the transparency and efficiency of their supply chains through the effective use of big data technology.

Keywords: Apple, supply chain transparency, principal component analysis, big data techniques.

1. Introduction

With the rapid advancement of big data technology in recent years, enterprises are gradually realising that through the collection, integration and analysis of massive data, they can achieve comprehensive monitoring and optimisation of the supply chain, thus enhancing transparency and reducing potential risks. However, supply chain data has high dimensionality, multiple indicators and complex correlations with each other, which brings great challenges to enterprises in data analysis. Principal Component Analysis (PCA) is a well-established statistical dimensionality reduction tool that can effectively handle high-dimensional data and simplify the analysis process by extracting key features.

This research leverages principal component analysis to extract the core features of suppliers and analyzes the financial data of some suppliers of Apple, through standardisation of data and dimensionality reduction analysis. Also, this paper aims to further verify the practicality of principal component analysis in processing complex data in the supply chain, and combines it with big data technology to provide a new analysis paradigm for the academic community. This research provides Apple and other global enterprises with quantitative methods to improve the transparency of the supply chain. and a reference for the digital transformation of enterprises in supply chain management, which is of great practical significance.

DOI: 10.54254/2754-1169/2025.19806

2. Literature Review

Recent studies have explored the role of PCA and big data in enhancing supply chain management. Sanders found that big data-driven supply chain management can help enterprises better optimise resource allocation in the face of market fluctuations, thus improving their operational efficiency through the study of several enterprise cases [1]. Meanwhile, Beattie et al. further proposed that principal component analysis has significant advantages in supply chain data feature extraction and simplifying complex models [2]. By analysing the main features of supply chain data, enterprises can have a clearer understanding of the key factors in the supply chain, so as to develop more targeted management strategies. These studies provide the theoretical basis and empirical reference for the development of this paper. According to Bi et al. supply chain transparency not only improves management efficiency, but also enhances trust between upstream and downstream partners, thus reducing friction and conflict caused by information asymmetry [3]. Ben-Daya et al. pointed out the application of Internet of Things (IoT) devices and sensor technology can collect key data in the supply chain in real time, providing enterprises with dynamic and comprehensive information flow [4]. Wang et al. showed that PCA has significant advantages in handling high-dimensional data in supply chains [5]. Papadopoulos et al. investigated the ability of high-transparency supply chains to cope with emergencies, such as natural disasters, and found that transparent supply chains are able to significantly reduce the impacts of supply chain disruptions on firms through data sharing and rapid decision-making [6].

Recent studies highlight the significant role of principal component analysis (PCA) and big data in enhancing supply chain management by optimizing resource allocation, improving transparency, and enabling rapid decision-making to effectively respond to market fluctuations and disruptions.

3. Methodology

In modern supply chain management, enterprises need to quickly and accurately identify key suppliers and potential risk points [7]. Apple, as a global leader in technology enterprises, has a complex and wide coverage supply chain system, and the performance of suppliers directly affects the productivity and market competitiveness of Apple products. For this reason, this paper uses Principal Component Analysis (PCA) to perform dimensionality reduction analysis on the financial data of some of Apple's suppliers, aiming at extracting the key characteristics of the suppliers to help Apple optimise its supply chain management.

3.1. Principles of Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique that maps the original high-dimensional data into a new low-dimensional space through a linear transformation, allowing the data in the low-dimensional space to retain as much of the original information as possible [8]. Its core idea is to calculate the covariance matrix between the characteristics of the original data, which is used to measure the correlation between the features. By decomposing the eigenvalue of the covariance matrix, the eigenvalue (representing the magnitude of the variance) and the corresponding eigenvector (the direction of the principal component) are obtained. The eigenvalues are then sorted in descending order, and the principal components that collectively account for at least 85% of the cumulative variance are selected. The original data is projected into a new space defined by these principal components, effectively reducing its dimensionality. The goal of PCA is to retain the main information of the data, simplify the analysis process, and reduce the complexity of dimensions.

DOI: 10.54254/2754-1169/2025.19806

3.2. Data Collection

Data acquisition and integration mainly include the use of suppliers' operating income, year-on-year growth rate and other data to build a core index system for supply chain management. Data standardisation and dimension reduction include extracting the main component through PCA, simplifying complex high-dimensional data into two key variables (principal component 1 and principal component 2, and visualising them.

The data in this paper includes operating revenues and year-over-year growth rates for some of Apple's suppliers, covering the following variables:

Operating income for the first half of 2020 ($bn)

Operating income for the first half of 2019 ($bn)

Year-on-year growth rate (%)

3.3. Data Processing

The following R language code [9] was used to implement the PCA analysis and visualisation of the results:

```
# Install and load the required packages
install.packages("ggplot2") # for visualisation
install.packages("stats") # for PCA
install.packages("scales") # for data normalisation
library(ggplot2)
library(stats)
library(scales)
# Data input
data <- data.frame(
  stock code = c('002241', '600745', '300207', '603986', '300433', '002600', '000049', '002655',
                 '000823', '300032', '002957', '300709', '002055', '603626', '002426', '600525'.
                 '000050', '601138', '300183'),
  First half of 2020 = c(155.73, 241.18, 115.1, 16.58, 155.68, 119.44, 75.74, 4.91, 22.97, 6.21, 7.93, 6.08,
                   30.55, 13.65, 42.76, 23.61, 140.57, 1766.54, 32.32).
  First half of 2019 = c(135.75, 114.34, 108.57, 12.02, 113.59, 95.96, 73.28, 4.57, 22.4, 7.79, 8.29, 5.47,
                  32.56, 8.95, 68.7, 28.96, 145.95, 1705.08, 32.32).
  Year-on-year growth rate = c(14.71, 110.93, 6.02, 37.91, 37.05, 24.47, 3.35, 7.3, 2.56, -20.31, -4.36, 11.12, -6.17, -4.36)
                 52.5, -37.76, -18.48, -3.69, 3.6, 0.0)
)
# Extract features
X <- data[, c("first half of 2020", "first half of 2019", "year-on-year growth rate")]
# Data standardisation
X_scaled <- scale(X)
# PCA downgrades
pca_result <- prcomp(X_scaled, centre = TRUE, scale. = TRUE)
# Getting principal component scores
df_pca <- data.frame(pca_result$x)
# Merge principal component scores into raw data
data$principal component1 <- df_pca$PC1
```

DOI: 10.54254/2754-1169/2025.19806

```
data$principal component2 <- df_pca$PC2
# Plotting PCA results
ggplot(data, aes(x = principal component 1, y = principal component 2)) +
  geom_point(colour = 'blue', alpha = 0.7) +
  geom_text(aes(label = ticker symbol), size = 3, vjust = -1) +
  labs(title = "Principal Component Analysis Results (PCA)", x = "Principal Component 1", y = "Principal Component 2") +
  theme_minimal() +
  theme(
    text = element_text(family = "Times New Roman"),
    axis.text = element_text(size = 10),
    plot.title = element_text(size = 14)
  )
```

4. Analysis of Results

Table 1: Statistics of principal component analysis results [10]

stock code (computing)	First half of 2020 ($bn)	First half of 2019 ($bn)	Year-on-year growth rate (per cent)	Principal Component 1	Principal component 2
002241	155.73	135.75	14.71	-0.01562837	0.09938775348129875
600745	241.18	114.34	110.93	0.12360802914696427	3.1783142785707565
300207	115.1	108.57	6.02	-0.144085261	-0.178048719
603986	16.58	12.02	37.91	-0.50120461	0.8424167284289563
300433	155.68	113.59	37.05	-0.052559006	0.8145811068411506
002600	119.44	95.96	24.47	-0.155700542	0.41233277894962966
000049	75.74	73.28	3.35	-0.284210919	-0.262437487
002655	4.91	4.57	7.3	-0.544180848	-0.134391369
000823	22.97	22.4	2.56	-0.478200557	-0.286183531
300032	6.21	7.79	-20.31	-0.542299175	-1.015637394
002957	7.93	8.29	-4.36	-0.534353922	-0.506617124
300709	6.08	5.47	11.12	-0.53939998	-0.012508057
002055	30.55	32.56	-6.17	-0.44703028	-0.565260106
603626	13.65	8.95	52.5	-0.508923697	1.3080728090278073
002426	42.76	68.7	-37.76	-0.363369371	-1.576186129
600525	23.61	28.96	-18.48	-0.469588924	-0.958193317
000050	140.57	145.95	-3.69	-0.028456738	-0.489951234
601138	1766.54	1705.08	3.6	5.928338440736221	-0.301465822
300183	32.32	32.32	0	-0.442754269	-0.368225166

According to Table 1, the following dot chart can be obtained.

DOI: 10.54254/2754-1169/2025.19806

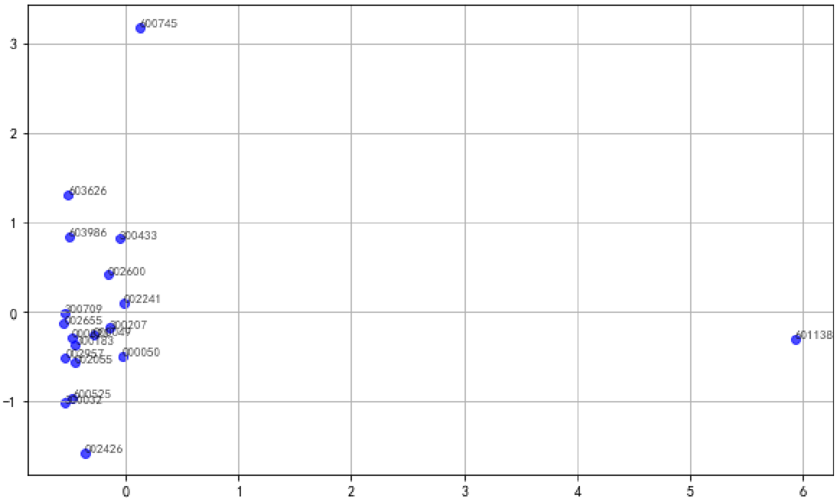

Figure 1: Distribution of principal component analysis results

PCA extracted two sets of principal components. Principal Component 1 mainly reflects the size characteristics of suppliers' operating revenues and explains most of the variance in the data. Principal component 2 mainly reflects the volatility of supplier growth rates and explains the dynamic trends of suppliers. The PCA downscaled data allows for a more concise understanding of the core characteristics of the supplier.

High-performing suppliers (high Principal Component 1), such as "Wintech (600745)" and "Industrial Fortune (601138)", are core suppliers in Apple's supply chain with outstanding performance in terms of operating revenue and growth rate. High-risk suppliers (with low or negative principal component 2), such as "Victory Precision (002426)" and "Golden Dragon (300032)", have large fluctuations in growth rates and may be exposed to operational risks. Stable suppliers like "Desai Battery (000049)" is a low-risk supplier with a stable growth rate despite its small revenue.

Figure 1 provides a clear view of the distribution position of each supplier across the principal component 1 and principal component 2 dimensions. This helps Apple quickly assess the positioning and risk of suppliers in the supply chain.

5. Discussion

Through principal component analysis (PCA), this paper successfully simplifies the complex data of Apple's suppliers and extracts key features. The results of the study show that PCA can effectively identify the performance differences of suppliers, which provides a quantitative basis for Apple to optimise its supply chain. Supply chain transparency is one of the core objectives in modern supply chain management, the significance of which is to help enterprises manage the complex supply chain network more efficiently and enhance their control over supply chain nodes [11]. Combined with visual analysis, enterprises can more intuitively assess the advantages and risks of suppliers in the supply chain, providing support for resource allocation and decision-making.

6. Conclusion

This study provides a new idea and research paradigm in the field of supply chain management, through the combination of big data technology and principal component analysis, to perform downscaling and feature extraction on the complex data of supply chain .The analysis results show that principal component 1 mainly reflects the scale characteristics of suppliers, while principal component 2 reveals the dynamic change characteristics of suppliers. Combining these results, this

DOI: 10.54254/2754-1169/2025.19806

paper provides a scientific basis for Apple to identify key suppliers, optimise resource allocation and improve risk identification in supply chain management.

Although this paper provides new perspectives and analyses in the field of supply chain management, there are still some shortcomings, which limit the breadth and depth of the research to a certain extent, and provide ideas for future directions of improvement. The data in this study is derived from the financial performance of some of Apple's suppliers, mainly covering financial indicators such as operating revenue and growth rate. However, supply chain transparency is a multi-dimensional issue that involves data on production capacity, logistics efficiency, environmental compliance and supply chain resilience in addition to financial data. Future research can integrate data from more dimensions, such as introducing indicators such as transport time of logistics nodes, suppliers' environmental emission data or order delivery accuracy, to build a more comprehensive supply chain management evaluation system.

References

[1] *Sanders, N. R. (2014). Big data driven supply chain management: a framework for implementing analytics and turning information into intelligence. Pearson Education.*

[2] *Beattie, J. R., Esmonde-White, F. W. (2021). Exploration of principal component analysis: deriving principal component analysis visually using Applied Spectroscopy, 75(4), 361-375.*

[3] *Bi, G. B., Ye, W., Xu, Y. (2024). The impact of information transparency on trade credit: the mediation role of risk. Kybernetes, 53(1), 27-57.*

[4] *Ben-Daya, M., Hassini, E., Bahroun, Z. (2019). Internet of things and supply chain management: a literature review. International Journal of Production Research, 57(15-16), 4719-4742.*

[5] *Wang, J., Swartz, C. L., Huang, K. (2023). Data-driven supply chain monitoring using canonical variate analysis. Computers & Chemical Engineering, 174, 108228.*

[6] *Papadopoulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S. J., Fosso-Wamba, S. (2017). The role of big data in explaining disaster resilience in supply chains for sustainability. Journal of Cleaner Production, 142, 1108-1118.*

[7] *Hugos, M. H. (2024). Essentials of supply chain management. John Wiley & Sons.*

[8] *Hasan, B. M. S., Abdulazeez, A. M. (2021). A review of principal component analysis algorithm for dimensionality reduction. Journal of Soft Computing and Journal of Soft Computing and Data Mining, 2(1), 20-30.*

[9] *Ihaka, R., Gentleman, R. (1996). R: a language for data analysis and graphics. Journal of computational and graphical statistics, 5(3), 299-314.*

[10] *Kaiwen, R. (2023). Research on the joint movement of stock public opinion and stock price based on text mining takes Apple's supply chain as an example. Nanjing University of Posts and Telecommunications*

[11] *Lamming, R. C., Caldwell, N., Phillips, W. E. (2004). Supply chain transparency. New et al, 191-208.*