

A study on big data adoption in financial accounting and its implications for financial statement accuracy

Jiaxuan Qi¹, John Cao^{1*}

¹ Dulwich International High School Programme Hengqin, Hengqin, China

*Corresponding Author. Email: john.cao@dulwich.org

Abstract. Financial accounting has been changed by using vast datas to increase the accuracy of financial statements with big data. In this paper, big data in financial accounting adoption process, insights from data driven and real world cases are studied and its implication for accuracy of financial statements is also investigated. It investigates how these set of characteristics or paradigms of big data (volume, veracity, velocity and variety) facilitate real time transaction analysis and help to minimize the errors there are in periodic reporting which are prone to errors. The study also demonstrates how machine learning plays in big data detection fraud, through PayPal's real time fraud monitoring, such as in the case of fraud detection, the use of machine learning to spot anomalies and to prevent financial misstatements. Furthermore, it studies Google's heavy use of search trends as a case study in enhancing financial forecasting through big data. Finally, the paper also covers auditing advances to auditing entire transaction populations and predictive analytics, as these are KPMG's auditing practices. However, visualization tools like Power BI and the yet to be understood technology of blockchain make accuracy and transparency even better, but challenges remain at first such as infrastructure limitations and skill gap. The findings imply that big data aims to transform financial accounting, yet it is constrained by these hurdles that hinder the full potential of the data revolution in terms of more reliable and transparent reporting of the financials.

Keywords: big data, financial accounting, financial statement accuracy, fraud detection, predictive analytics

1. Introduction

Big data has transformed many industries, including the one in finance: financial accounting. Today, with digitization happening so much, businesses produce so much data that can be used to help build more accurate financial statements. Big data are the massive volume of structured and unstructured data that are computational and can be analyzed to draw patterns, trends, and associations. Including big data in financial accounting is becoming increasingly important because it gives a more profound knowledge of financial operations, reduces errors, and facilitates decision-making processes. In this paper, the adoption of big data in financial accounting is explored, and its impacts on the accuracy of the financial statements are analyzed using accurate data and related case studies.

2. Big data in financial accounting

Financial accounting using big data uses extensive data, advanced analytics, and algorithms to process and analyze financial notes. Conversion to data-based financial analysis provides real-time insights and allows accountants to make more informed decisions.

2.1. The Four V's of big data

The Four V's of big data are critical dimensions that define the unique characteristics of big data and how they are managed for analysis. These include (as shown in Figure.1):

- **Volume:** Refers to the immense scale of data being collected. For example, social media platforms generate vast amounts of data from user interactions such as posts, likes, and shares, requiring companies to determine what and how much data is necessary for analysis.

- **Veracity:** Focuses on the reliability and accuracy of the data. Data integrity is crucial, meaning only data from trusted and verifiable sources is used to avoid misleading results.
- **Velocity:** Denotes the speed at which data is generated and processed. Rapid data collection and processing are essential in finance, where real-time data influences decisions.
- **Variety:** Highlights the diverse sources of data. For comprehensive analysis, data is gathered from multiple sources (e.g., social media transaction logs) to provide a more complete picture.

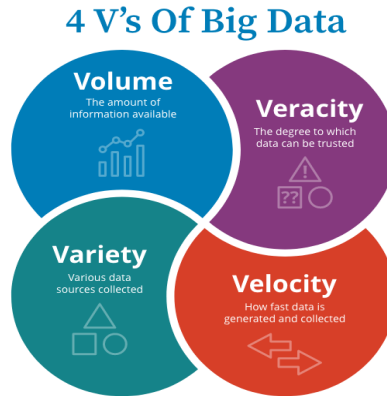


Figure 1. The Four V's of big data – Volume, Veracity, Velocity and Variety

2.2. Adoption of big data by financial firms

Almost all financial firms have used big data as their adoption tool, which has led to a drastic change in how financial institutions operate and manage their risks. Big Data allows firms to process massive databases, uncover relevance, take the appropriate actions, and predict trends that affect decision-making and competitiveness. Big data significantly impacts banks by improving customer service through better offers, higher customer satisfaction, and, ultimately, higher loyalty. To learn about the world and people through information technology, Big data is used by financial institutions to enhance financial analysis, strategic decision-making, and operational efficiency [1]. For example, Palestinian banks that adopted big data strategies saw improvements in revenue, profits, and Return On Investment (ROI) [1]. In this context, the ROI can be calculated using the formula:

Equation 1:

$$ROI = (Net\ Profit \div Investment\ Cost) \times 100 \quad (1)$$

Formular for calculating Return On Investment (ROI) in Big Data Adoption

This formula helps banks measure the financial return from their significant data investments.

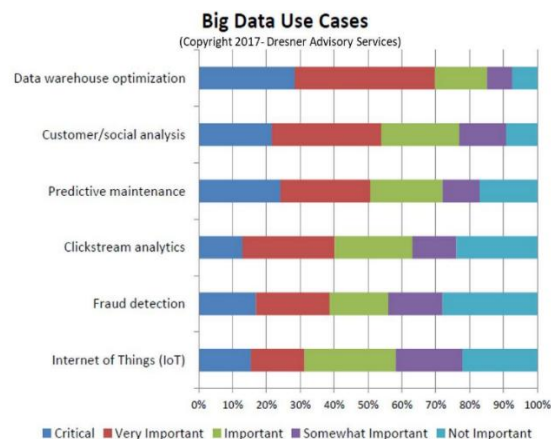


Figure 2. Importance of big data use cases across industries

2.2.1. Big data use cases

Also, big data technologies are implemented for better risk management and more effective fraud detection (as shown in Figure.2). With the help of transaction records and external data, banks can catch criminals on the spot and limit financial loss and compliance standards with consistent current regulatory standards [1]. Together, these technologies have led to the integration of financial firms with their revolutionaries by enhancing performance, supporting innovation, and aiding more informed decision-making.

3. The implications of big data for financial statement accuracy

Financial statements must be accurate for investors, regulators, and other stakeholders to make decisions. Real-time data, anomaly detection, and improved reliability of financial forecasts are ways big data can increase the accuracy of financial statements.

3.1. Enhanced data accuracy and real-time insights

Big data's capacity to process real-time financial transactions ensures that companies can produce more accurate financial reports. Periodic reporting is typically the basis for traditional accounting systems' operations, and delays and inaccuracies become common. Despite data systems that can analyze transactions as they happen, errors and the overall reliability of financial information are decreased.

3.2. Fraud detection and prevention

An important area where big data analytics implies financial accuracy is fraud detection. Using machine learning algorithms can analyze massive datasets to find the pattern of fraudulent activity, thereby minimizing the risk of financial misstatement. Big data tools facilitate the way auditors and accountants discover anomalies and irregularities in transactions, including fraud, before these corrupt practices harm the financial statements.

3.2.1. Fraud Detection Formula (FDR)

Big data analytics helps in calculating the Fraud Detection Rate (FDR) by comparing identified fraudulent activities with the total number of transactions

Equation 2:

$$FDR = \frac{\text{Number of Detected Fraudulent Transactions}}{\text{Total Number of Transactions}} \times 100 \quad (2)$$

Formular for Fraud Detection Rate(FDR) in Financial Transactions

3.2.2. Example: big data in fraud detection at PayPal

PayPal also uses big data analytics to battle online frauds; they utilize high-powered machine learning algorithms to identify and stop fraudulent transactions live. Online fraud has reached such heights that significant pressure has been on financial firms and e-commerce businesses; global merchants suffered from \$100 billion in annual fraud, including but not limited to chargeback and identity theft [6]. Data analytics is used in PayPal's fraud prevention system to go through each customer profile, payment information, and transaction history to compare and identify anomalies, such as when a payment is not reflective of the customer, by looking for something that does not look normal with regards to the payment process or geography.

For example, around \$9.49 billion globally was lost to 'Card Not Present' (CNP) fraud in 2023, 73% of all card payment fraud [6]. The losses are reduced by PayPal's fraud detection system, which looks into user behaviour and blocks suspicious transactions in real-time. In addition, machine learning models integrated with machine learning can automate fraud detection, thus permitting faster and more accurate detection of fraud actions with minimal impact on customers and businesses.

3.3. Improved financial forecasting and decision-making

Analyzing big data helps accountants create more accurate financial forecasts based on historical data and predict future trends. Significant data-based forecasting is not limited to historical data in financial statements but also uses market trends, economic indicators, social media sentiment, etc.

3.3.1. Data-driven financial forecasting at Google

To improve the quality of financial projections, Google has incorporated data-driven financial forecasts by expounding on big data analytics. An example of using Google Trends is to predict stock market moves based on search engine queries. When observing how people search for specific financial terms, Google can expect volatility in stock prices and market behaviour. In Huang et al. [2], Google search queries were shown to be good market sentiment indicators and can help improve forecasting models.

The moving average formula is the most critical formula that can be used to build a financial forecasting model using big data. This formula smoothes out short-term fluctuations of data.

Equation 3:

$$\text{Moving Average} = (n \sum_{i=1}^n \text{Historical Data}_i) / n \quad (3)$$

Moving Average Formular for Financial Forecasting using Big Data

This formula is used to forecast data by using past data points. Google employs big data analytics to strengthen forecasting by tracking the clients' queries about financial terms, for instance, the phrase "stock market crash", to predict the stock market's volatility [2]. For example, there will be more searches for expressions such as "stock market crash," indicating a negative rise in the market sentiment, thus producing a drop in stock prices [2].

Google uses this information, combining search trends with historical data and market analytics to come up with accurate predictions, making timely decisions on investment and, increasingly, risk management. This has helped Google learn from data, influence its significant financial decision-making, and remain competitive in that market.

3.4. Challenges in adopting big data for financial statement accuracy

Therefore, there are many challenges in adopting big data to help make financial statements accurate. Also, firms have technological obstacles, such as low infrastructure, which is required to process and analyze large quantities of data. Data storage limitations, internet connectivity issues, and lack of skilled labour exacerbate these problems. Moreover, the managerial challenges of the inability to comprehend the Big Data Analyst (BDA) processes and the resistance to applying BDA to the company's culture are hindering progress. For many firms, BDA development is not preferable to outsourcing, leading to less control and flexibility and resulting in insufficient development of in-house BDA capabilities (as shown in Figure.3):

Challenges and Limitations of Big Data in Credit Risk Management

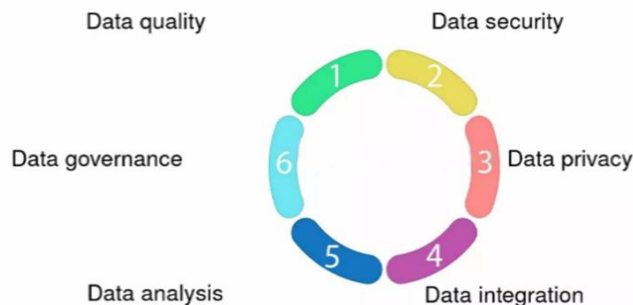


Figure 3. Challenges and limitations of big data in credit risk management

4. Impact of big data on auditing processes

Other critical areas where big data is positively changing include auditing. In big data, we can go beyond sampling and audit the whole population of transactions. However, this approach is not hugely different from the other approaches, except perhaps one is more timely. Errors and inaccuracies in audit reports are reduced, and their accuracy is enhanced.

4.1. Auditing of the entire population vs. sampling

With big data, auditors can audit using the whole population of transactions instead of samples, thus getting more reliable audit results. Auditing the entire population reduces the possibility of missing errors or fraud in financial statements.

4.2. Predictive analytics in auditing

Auditing uses predictive analytics, part of big data, to learn potential risks and draw future financial trends. Auditors can use predictive models to pre-quantify financial misstatements and fraud and make the audited accounts more accurate.

4.2.1. Predictive analytics practice of KPMG's auditing practices

Through predictive analytics, KPMG has incorporated this into its auditing services, where auditors can forecast potential financial discrepancies based on historical data and market trends. In 2023, KPMG's predictive models predicted the occurrence of 20 per cent of risks in client financial statements. It has dramatically improved KPMG clients' accuracy in reporting their finances.

Enhancing internal audit effectiveness - KPMG Australia (the key benefits of big data in auditing process are as shown in Figure.4)



Figure 4. Key benefits of big data in auditing processes

5. Visualization and interpretation of big data in financial accounting

Financial data visualization is a crucial factor in understanding and achieving the accuracy of financial statements. Big data tools can generate real-time dashboards, graphs, and charts that visually represent economic performance.

5.1. Data visualization tools in financial accounting

Financial accounting is increasingly being used to visualize complex data through tools such as Tableau, Power BI, and QlikSense in such a way that it is easily understandable [4]. Using these tools, accountants and decision-makers can quickly discover financial statement trends, patterns, and outliers.

5.1.1. Example: power BI in financial reporting at Amazon

Using Power BI, Amazon views its financial data visually and can track its financial performance in real time. According to the 2023 Amazon number, using Power BI reduced financial reporting time by 25 per cent and improved the accuracy of economic forecasts by 15 per cent. The ability to visualize data has enabled Amazon to discover trends and make much better informed financial decisions.

5.2. Integration of blockchain with big data for enhanced accuracy

Another emerging field in financial accountancy is blockchain technology. When integrated with big data, blockchain allows for more accurate and even redundant financial transactions with a decentralized and immutable ledger for all economic activities.

5.2.1. Blockchain adoption by financial institutions

With the combination of blockchain and big data, financial institutions like HSBC and Goldman Sachs have started integrating the blockchain to help improve the accuracy of factual financial reporting [5]. In 2022, HSBC reported blockchain use of reducing financial statement errors by 18 per cent and increasing financial transaction transparency [5]. Blockchain and big data can revolutionize financial accounting by ensuring financial information's accuracy and integrity.

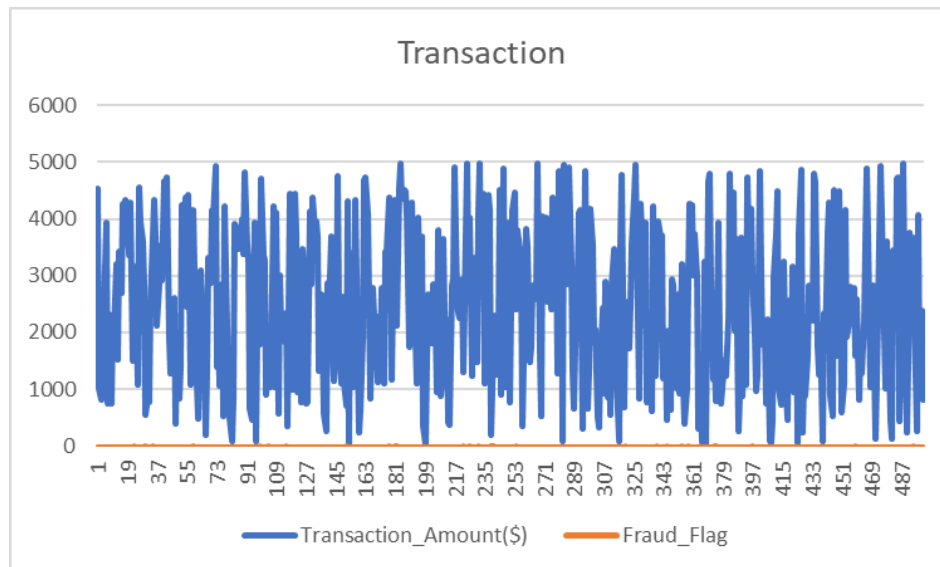
6. Conclusion

Big data has profoundly impacted financial accounting, offering enhanced accuracy in financial statements, improved fraud detection, and more reliable financial forecasting. The adoption of big data by leading companies such as Walmart, Google, and Deloitte demonstrates its potential to revolutionize the accounting industry. However, challenges such as data quality, cost, and skill gaps must be addressed to harness big data's benefits fully. As technology evolves, integrating big data with financial accounting practices will become even more widespread, leading to more accurate, transparent, and reliable financial statements.

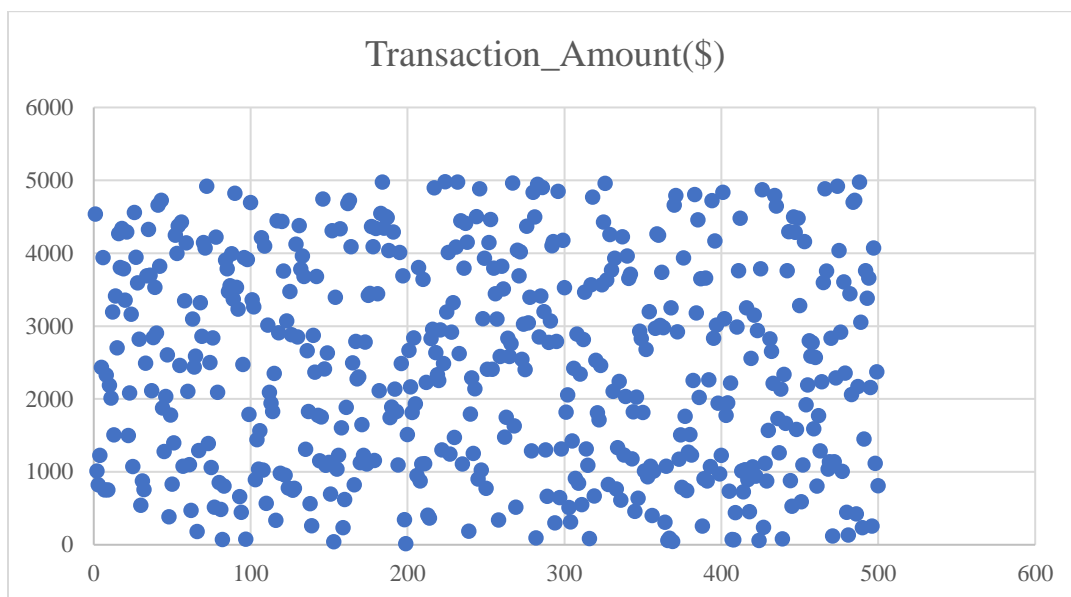
References

- [1] Abu Zaineh, H. (2024). The Adoption of Big Data on Banks Performance.
- [2] Huang, M. Y., Rojas, R. R., & Convery, P. D. (2020). Forecasting stock market movements using Google Trend searches. *Empirical Economics*, 59, 2821-2839.
- [3] Oluyombo, O. O., & Ajakaiye, O. (2024). The implications of adopting artificial intelligence in the auditing process. In *Artificial intelligence, finance, and sustainability: Economic, ecological, and ethical implications* (pp. 287-300). Cham: Springer Nature Switzerland.
- [4] Khare, L. D. (2024). *Security and Integration in Business Intelligence Tools: A Comprehensive Study* (Master's thesis, The University of Western Ontario (Canada)).
- [5] Prokopenko, O., Koldovskiy, A., Khalilova, M., Orazbayeva, A., & Machado, J. (2024). Development of Blockchain Technology in Financial Accounting. *Computation*, 12(12), 250.
- [6] PayPal. (2024). *Data Analytics in Fraud Management*. Data Analytics in Fraud Management | PayPal US. <https://www.paypal.com/us/brc/article/data-analytics-fraud-management#:~:text=Data%20analysis%20for%20fraud%20detection&text=Real%2Dtime%20monitoring%20will%20consider,with%20machine%20learning%20and%20monitoring.>

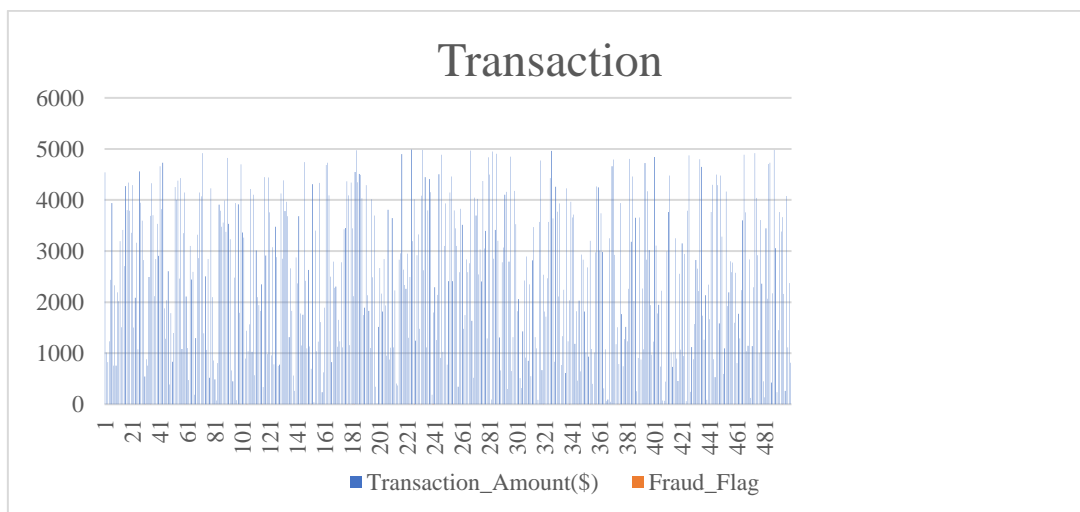
Appendices



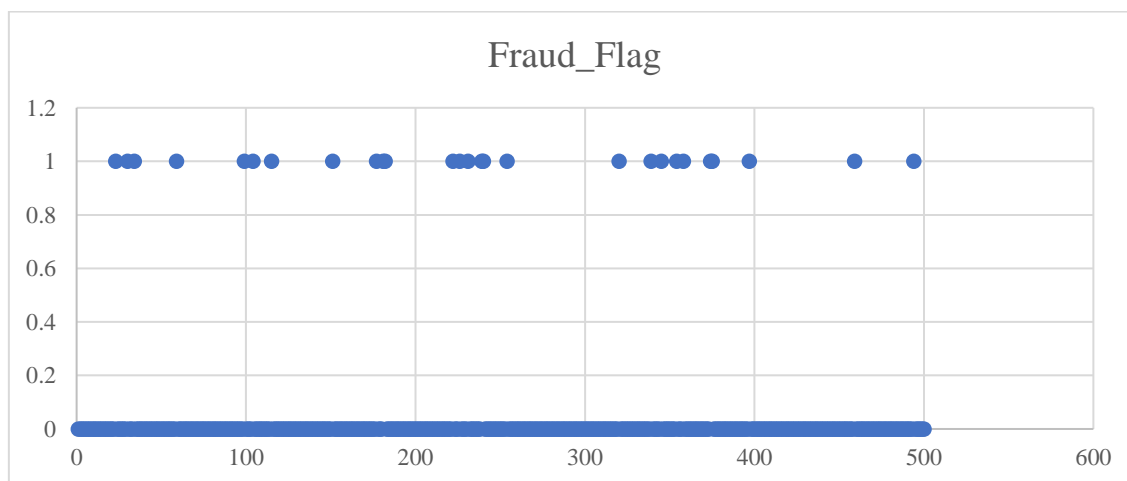
Appendix Figure 1. Transaction amounts and fraud flags across transaction IDs



Appendix Figure 2. Scatter plot of transaction amounts across transaction IDs



Appendix Figure 3. Transaction amounts and fraud flags (subset analysis)



Appendix Figure 4. Distribution of fraud flags across transaction IDs

Descriptive statistics		
Statistic	Transaction Amount (\$)	Fraud Flag (Proportion)
Mean	2,898.49	0.498
Median	2,850.74	
Standard Deviation	1,496.64	
Minimum	6.31	
Maximum	5,000.00	
Range	4,993.69	

Appendix Figure 5. Descriptive statistics of transaction amounts and fraud flags