International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 6 Issue 3, March-April 2022 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

# Big Data in Economics: An Introduction

**Matthew N. O. Sadiku[1], Uwakwe C. Chukwu[2], Abayomi Ajayi-Majebi[3], Sarhan M. Musa[1]**

[1]Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, USA
[2]Department of Engineering Technology, South Carolina State University, Orangeburg, SC, USA
[3]Department of Manufacturing Engineering, Central State University, Wilberforce, OH, USA

## ABSTRACT
Big data is a domain which is growing rapidly in every sector. In just a few years, big data has impacted industries and activities from marketing to law enforcement. It is often regarded by optimists as a revolution that will change for the better how we live, think, and work. Big data has a great impact on the economy since it helps companies, organizations, and governments to efficiently manage large volumes of data. This paper addresses the impact of big data in the economic field and provides some applications of big data in economy.

***KEYWORDS:*** *big data, economics, economists, applications*

***How to cite this paper:*** Matthew N. O. Sadiku | Uwakwe C. Chukwu | Abayomi Ajayi-Majebi | Sarhan M. Musa "Big Data in Economics: An Introduction" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-6 | Issue-3, April 2022, pp.1749-1755, URL: www.ijtsrd.com/papers/ijtsrd49839.pdf

IJTSRD49839





## INTRODUCTION
Victor Hugo once said, "No power on earth can stop an idea whose time has come." Big data is one such wonderful idea. Data is a valuable commodity. As shown in Figure 1, data is everywhere in an interconnected world [1]. The role of economists in technology companies is growing due to the ubiquity of data. The quality and quantity of data on economic activity are growing fast. Today, economists are increasingly making use of newly available massive administrative data or private sector data to create new opportunities. They are also increasingly adopting the large-data statistical methods to complement traditional econometric techniques [2].

Big data refers to data sets of much larger size and higher frequency. It may be regarded as an untapped resource that technology finally allows us to exploit. It usually consists of large sets of unstructured, semi-structured, or structured data obtained from numerous sources. Big data comes from various sources, as shown in Figure 2 [3]. It is data generated by the increasing use of digital devices such digital sensors, smart meters, satellite imagery, administrative records, medical records, weather data, business transaction systems, social networks, mobile applications, scientific experiments. etc Data obtained from non-traditional socio-economic sources are generally large, heterogeneous and unstructured or semi-structured. The big data era creates a lot of exciting opportunities for new developments in economics.

## REVIEW ON BIG DATA
Big data (BD) refers to a collection of data that cannot be captured, managed, and processed by conventional software tools. It is a relatively new technology that can help many industries. The three main sources of big data are machines, people, and companies. Big data can be described with 42 Vs [4]. The first five Vs are volume, velocity, variety, veracity, and value [5].

- **Volume:** This refers to the size of the data being generated both inside and outside organizations

and is increasing annually. Some regard big data as data over one petabyte in volume.

- **Velocity:** This depicts the unprecedented speed at which data are generated by Internet users, mobile users, social media, etc. Data are generated and processed in a fast way to extract useful, relevant information. Big data could be analyzed in real time, and it has movement and velocity.
- **Variety:** This refers to the data types since big data may originate from heterogeneous sources and is in different formats (e.g., videos, images, audio, text, logs). BD comprises of structured, semi-structured or unstructured data.
- **Veracity:** By this, we mean the truthfulness of data, i.e. weather the data comes from a reputable, trustworthy, authentic, and accountable source. It suggests the inconsistency in the quality of different sources of big data. The data may not be 100% correct.
- **Value:** This is the most important aspect of the big data. It is the desired outcome of big data processing. It refers to the process of discovering hidden values from large datasets. It denotes the value derived from the analysis of the existing data. If one cannot extract some business value from the data, there is no use managing and storing it.

On this basis, small data can be regarded as having low volume, low velocity, low variety, low veracity, and low value. Additional five Vs has been added [7]:

- **Validity:** This refers to the accuracy and correctness of data. It also indicates how up to date it is.
- **Viability:** This identifies the relevancy of data for each use case. Relevancy of data is required to maintain the desired and accurate outcome through analytical and predictive measures.
- **Volatility:** Since data are generated and change at a rapid rate, volatility determines how quickly data change.
- **Vulnerability:** The vulnerability of data is essential because privacy and security are of utmost importance for personal data.
- **Visualization:** Data needs to be presented unambiguously and attractively to the user. Proper visualization of large and complex clinical reports helps in finding valuable insights.

For example, Figure 3 shows the 8V's of big data [7]. In addition, the 10V's above, some suggest the following 5V's: Venue, Variability, Vocabulary, Vagueness, and Validity) [8].

To thrive in today's complex business environment, businesses must adopt a data-driven culture and leverage analytics platforms to make key decisions that improve productivity. Industries that benefit from big data include the healthcare, financial, airline, travel, restaurants, automobile, sports, agriculture, manufacturing, and hospitality industries. The different stages of a full data lifecycle in the context of economic and social analyses are illustrated in Figure 4 [9].

## BIG DATA ANALYTICS

Every day, data is growing bigger and bigger, and big data analysis (BDA) has become a requirement for gaining invaluable insights into data such that companies could gain significant profits in the global market. Once the big data is ready for analysis, we use advanced software programs such as Hadoop, MapReduce, MongoDB, Spark, Cassandra, Apache Storm, and NoSQL databases [10]. Big data analytics refers to how we can extract, validate, translate, and utilize big data as a new currency of information transactions. It is an emerging field that is aimed at creating empirical predictions. Data-driven organizations use analytics to guide decisions at all levels [11].

Here are some ways to effectively handle big data [7]:

1. **Outline Your Goal:** It is of paramount importance that organizations should collect data with a laser focus to benefit business objectives.
2. **Secure the Data:** We must ensure the relevant data collected is secured with a broad range of measures.
3. **Keep the Data Protected:** Organizations must safeguard databases against adverse environmental situations which would damage data and put forth considerable efforts to protect their data.
4. **Do Not Ignore Audit Regulations:** Irrespective of the data nature being payment data, credit scores or data of lesser importance, the data should be managed accordingly.
5. **Data Has to Be Interlinked:** All organizational data must be able to talk to each other.
6. **Know the Data You Need to Capture:** Organizations are required to know which data has to be collected and also when.
7. **Adapt to the New Changes:** Software and data in all its forms change constantly and almost on a daily basis, globally and one must be flexible.

## APPLICATIONS

We live in a data revolution and this has important implications for the field of economics. Some

researchers and media believe that big data will transform business, government, and other aspects of the economy. Big data is increasing in importance as a source of information about the economic world. The success of modern economics largely depends on the availability of data sources which allowed us to quantify human behavior. Common applications of big data in economics include the following:

- **Economic Analysis:** The application of big data in economic analysis could be associated with two notions [12]: (a) "Multidimensionality": in terms of number of variables per observation, number of observations or both. (b) "Granularity": Big data sets often provide useful micro-level data for analyzing agent behavior. traditional survey data. The high level of granularity in the data makes it possible to link the individual scores to the school tests and the corresponding tax records for a large sample, which would have been impossible with aggregate data or a smaller sample
- **Economic Statistics:** This describe the state of an economy. They span several sectors such as business statistics, macroeconomic statistics, economic accounts, and energy, industry, financial, tourism, and labor and trade statistics. Traditionally, much economic data are collected through surveys of individuals, households, and administrative systems. Big data sources can be used to produce economic statistics using online price data, scanner data, mobile phone data, Earth Observations, financial transactions data and smart meter data. There is a multitude of big data projects in the area of economic statistics undertaken by various nations [13].
- **Predictive Modeling:** Recent advances in the fields of statistics, machine learning, and econometrics have generated interest in economic predictive models and many predictors. The most common uses of big data by companies are for tracking business processes and for building a wide array of predictive models. For example, Amazon and Netflix recommendations rely on predictive models of what products a customer might want to purchase. Google's search relies on algorithms that predict the relevance of particular web pages. Online advertising and marketing rely heavily on automated predictive models that target customers. One Palo Alto company, Palantir, develops algorithms to identify terrorist threats and to detect fraudulent behavior in health care and financial services [14].
- **Financial engineering:** This is the process of creating innovative solutions for the existing company's financial problems using mathematical methods. Financial engineering uses tools and knowledge from the fields of computer science, big data, data science, data analytics, statistics, economics, and applied mathematics to address present financial issues and innovative financial products. It is helpful in derivative pricing, financial regulation, execution, corporate finance, portfolio management, risk management, etc. For example, financial engineering is used by Commercial Banks, Investment Banks, Insurance companies, and other fund hedging agencies. Big data, data science, and data analytics in financial engineering can be used as a successful tool at all stages of insurance business management practices [15].
- **Behavioral Economics:** Behavioral economics may be regarded as economic analysis that applies psychological insights into human behavior to explain economic decision-making. Much of our daily life is influenced by "big data" and "behavioral economics." Pricing and comparison strategies are meant to influence our decision-making in such a way to facilitate buying a particular product at a particular time to meet our perceived need. The coronavirus disease 2019 (COVID-19) pandemic has placed a magnifying glass on how individuals in our society make decisions that affect themselves, their immediate family, and the general public. Both big data and behavioral economics have played an enormous role in how pandemic care is implemented across the country [16].

## BENEFITS

The rise of big data is an exciting time with opportunities for economists, policymakers, or social scientists. Never before has so much data been available to test existing theories and develop new ones. The benefits of using big data for economic analysis are numerous. Big data allows for better prediction of economic phenomena and improves causal inference. Economists have thus moved from being forecasters to now casters.

More data potentially available for research is better than less data.

Big data is having a significant impact on our knowledge of the world, which plays a vital role in increasing economic activity and living standards. For example, we are often excited about data on weather, stock performance, and crop plantings. It is now possible to cost-effectively collect these data, use them in an informed manner, and turn them into actionable knowledge. The effective use of data has the capacity to transform every industry.

In addition, these are the ways big data will affect the economy [17,18]:

- Producing new goods and services, such as the Nest home thermometer or mass customized shoes.
- Big data allows for better prediction of economic phenomena and improves causal inference.
- More-targeted marketing that injects customer feedback into product design.
- Better organizational management.
- Faster innovation through a shorter research and development cycle.
- Errors inside the business are known immediately.
- The plan of action of your opposition is seen promptly.
- Extortion can be recognized the minute it happens and legitimate measures can be taken to restrict the harm.
- The principal points of interest of big data include the increased speed, capacity, and scalability of storage and having the measures and tools to deal with the data all the more proficiently.
- Big data technologies help companies store large volumes of data while enabling significant cost benefits.
- Improves law enforcement and security.
- Creates new business opportunities.
- Contributes to the living standards of people around the world and reduces poverty.

## CHALLENGES

The unprecedented rate at which high-dimensional individualized data are being generated means that big data poses a host of challenges. Data opportunities raise some important challenges. A major challenge of big data is managing large-scale data with many variables. As the amount of available data increases, all methods will tend to improve in terms of their predictive accuracy. New sources of data often create challenges that may require new skills [19].

In spite of the enormous potential of big data, integrating and analyzing the wide variety of heterogeneous sources cannot be tackled with the traditional methods used in economics. Predictions based on big data may have privacy concerns. Some methods should be developed for researchers to access and explore data in ways that respect privacy and confidentiality concerns. There is the common perception that big data tools focus exclusively on prediction at the expense of causal inference.

Other challenges of big data economics include [3]:

- A major challenge for economists is the scale of modern data sets. The sheer size of the data involved may require more powerful data manipulation tools.
- Big data is costly to collect, store, and analyze because it requires investments in technology and human skill.
- The data collected is raw, inconsistent, and therefore subjected to more noise.
- Big data struggles with security issue, especially on the social media front.
- Pinpointing data that flows into an organization on a daily basis is a lot more challenging than finding the proverbial needle in the haystack.
- Data selection is bias.
- New data sets may change the statistical methods used by economists.
- Lack of skilled data scientists.
- Analyses based on big data will focus too much on correlation and prediction.
- Big data can create a new digital divide.

In spite of these challenges, the next few decades are likely to be a very exciting time for economic research.

## TRAINING FUTURE ECONOMISTS

Economists mostly deal with the problems of two types. First, they figure out how *n* big variables, like inflation and unemployment, interact with each other. Second, they make practical policy recommendations for the business executive. Economists have traditionally dealt with data that fits in a spreadsheet, but that is changing fast as new more-detailed data becomes available.

It is crucial to train the next generation of researchers, policy leaders, and economists on methods to study and improve economic opportunities. Such training should complement traditional economics courses. Topics may include equality of opportunity, education, health, the environment, and criminal justice [20].

## CONCLUSION

Due to technological advances, data will become even easier to collect, transmit, store, and analyze. In the recent past, big data has experienced significant growth because institutions, organizations, and

businesses are heavily utilizing big data to achieve their objectives. It is significantly affecting the global economy. Since the growth in data production is highly unlikely to abate, the big data bubble will continue in the near future. The advantages of using big data for economic analysis are legion. More information about big data in economics can be found in the books in [21-28] and the following related journals:

- Journal of Big Data
- Journal of Economic Perspectives
- International Journal of Social Science add Economic Research

**REFERENCES**


[1] R. Krishanmurthy, "Big data in economics and policy," October 2020, https://www.analyticssteps.com/blogs/big-data-economics-and-policy

[2] L. Einav and J. Levin, "Economics in the age of big data," Science, vol. 346, no. 6210, November 2014.

[3] "The world of big data," August 2019, https://www.geeksforgeeks.org/world-big-data/

[4] "The 42 V's of big data and data science," https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html

[5] M. N. O. Sadiku, M. Tembely, and S. M. Musa, "Big data: An introduction for engineers," Journal of Scientific and Engineering Research, vol. 3, no. 2, 2016, pp. 106-108.

[6] P. K. D. Pramanik, S. Pal, and M. Mukhopadhyay, "Healthcare big data: A comprehensive overview," in N. Bouchemal (ed.), Intelligent Systems for Healthcare Management and Delivery. IGI Global, chapter 4, 2019, pp. 72-100.

[7] S. Josh, "What is the effective way to handle big data?" April 2017, https://www.zarantech.com/blog/effective-way-handle-big-data/

[8] J. Moorthy et al., "Big data: Prospects and challenges," The Journal for Decision Makers, vol. 40, no. 1, 2015, pp. 74–96. https://www.grandviewresearch.com/industry-analysis/industrial-wireless-sensor-networks-iwsn-market

[9] D. Blazquez and J. Domenech, "Big data sources and methods for social and economic analyses," Technological Forecasting and Social Change, vol. 130, May 2018, pp. 99-113.

[10] M. N. O. Sadiku, J. Foreman, and S. M. Musa, "Big data analytics: A primer," International Journal of Technologies and Management Research, vol. 5, no. 9, September 2018, pp. 44-49.

[11] C. M. M. Kotteti, M. N. O. Sadiku, and S. M. Musa, "Big data analytics," Invention Journal of Research Technology in Engineering & Management, vol. 2, no. 10, Oct. 2018, pp. 2455-3689.

[12] K. Balar and R. Chaabita, "Big data in economic analysis: Advantages and challenges," International Journal of Social Science add Economic Research, vol. 4, no. 7, July 2019, pp. 5196-5205.

[13] "Big data for economic statistics," Stats Brief, no. 28, March 2021.

[14] L. Einav and J. Levin, "The data revolution and economic analysis," Innovation Policy and the Economy, vol 14, 2014.

[15] V. Chakravaram et al., The role of big data, data science and data analytics in financial engineering," Proceedings of the 2019 International Conference, June 2019.

[16] R. Jhaveri, "Big data and behavioral economics in infectious diseases," Clinical Therapeutics, vol. 43, no. 10, 2021.

[17] J. Kennedy, " Big data's economic impact," https://www.ced.org/blog/entry/big-datas-economic-impact

[18] L. Robinson, "The central European review of economics and management," https://www.austriancenter.com/big-data-global-economy/

[19] "Big data in economics," https://wol.iza.org/articles/big-data-in-economics/long#:~:text=Big%20Data%20refers%20to%20data,to%20outperform%20more%20complex%20techniques

[20] R. Chetty, "Using big data to solve economic and social problems," https://opportunityinsights.org/course/

[21] L. Matyas et al. (eds.), Economics without Borders. Cambridge University Press, 2017.

[22] K. G. Abraham et al.. (eds.), Big Data for Twenty-First-Century Economic Statistics. University of Chicago Press, 2019.

[23] C. Tang, The Data Industry: The Business and Economics of Information and Big Data. Hoboken, NJ: Wiley, 2016.

[24] I. Aldridge and M. Avellaneda, Big Data Science in Finance. Wiley, 2021.

[25] S. Consoli, D. R. Recupero, and M. Saisana (eds.), Data Science for Economics and Finance: Methodologies and Applications. Springer; 2021.

[26] R. D. Francesco, Big Data Economics, Towards Data Market Places: Nature of Data, Exchange Mechanisms, Prices, Choices, Agents & Ecosystems. Independently Published, 2018.

[27] B. Schmarzo, The Economics of Data, Analytics, and Digital Transformation: The theorems, laws, and empowerments to guide your organization's digital transformation. Packt Publishing, 2020.

[28] A. M. Dima and M. Kelemen (eds.), Digitalization and Big Data for Resilience and Economic Intelligence: Proceedings of 4th International Conference on Economics and Social Sciences, ICESS 2021, Bucharest, Romania. Springer, 2022.


**Figure 1 Data is everywhere in an interconnected world [1].**

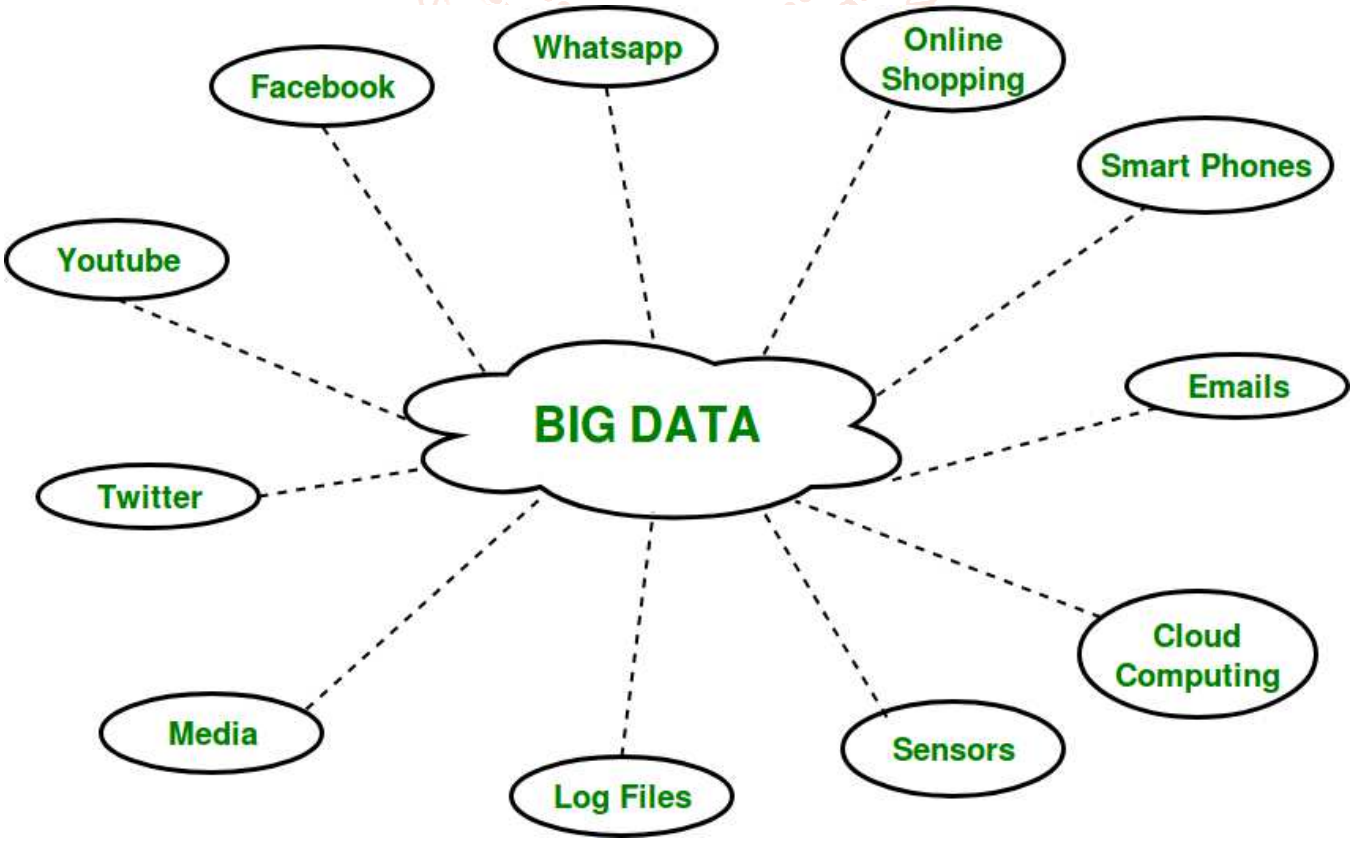


**Figure 1 Big data comes from various sources [3].**


**Figure 3 The 8V's of big data [7].**

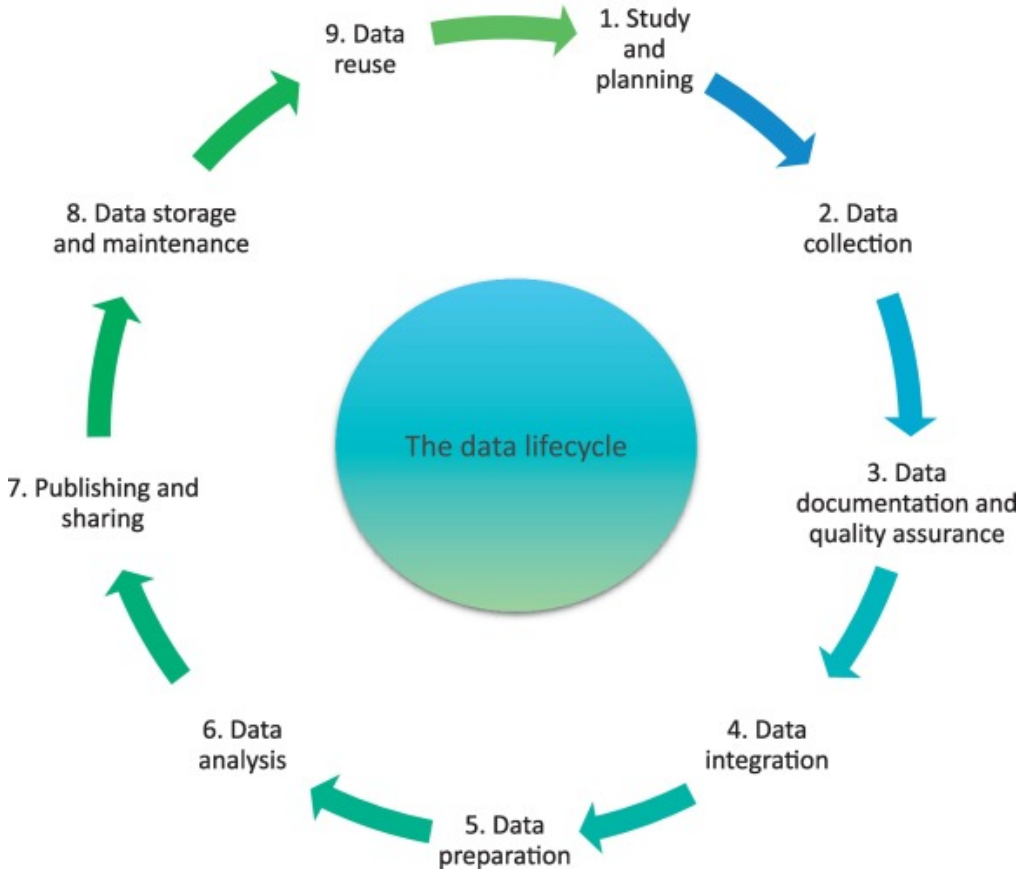


**Figure 4 The different stages of a full data lifecycle in the context of economic and social analyses [9].**