Learning Module 11: Introduction to Big Data Techniques

Q.3692 Alternative data refers to:

A. Data used for investment analysis arising from external sources, including financial statements and management presentations of comparable entities

B. Data used by investors to evaluate a company or product that is not related to financial statements

C. Data used by investors for investment analysis that is not within their traditional sources

The correct answer is **C.**

Alternative data, also known as non-traditional data, refer to data types generated by the use of electronic devices, social media, satellite and sensor networks, and company exhaust. Alternative data helps investors get more granular and faster insights into company performance compared to traditional data sources.

A is incorrect. This option describes data arising from external sources, including financial statements and management presentations of comparable entities. However, this description aligns more closely with traditional data sources rather than alternative data. Traditional data sources, such as financial statements and management presentations, have been the cornerstone of investment analysis for decades. These sources provide a historical view of a company's financial health and performance but may not offer the timely or unique insights that alternative data can provide.

B is incorrect. While this option touches on the essence of alternative data by mentioning data not related to financial statements, it does not fully capture the scope and utility of alternative data in investment analysis. Alternative data is not merely about being unrelated to financial statements; it is about providing additional, non-traditional insights that complement traditional analysis. This option fails to emphasize the broad range of data types and sources that constitute alternative data and how they are used by investors to gain a deeper understanding of market dynamics, consumer behavior, and company performance beyond what traditional financial metrics can reveal.

CFA Level I, Quantitative Methods, Learning Module 11: Introduction to Big Data Techniques. LOS (c): Describe applications of Big Data and Data Science to investment management.

Q.3700 The big data revolution witnessed in the last 50 years is down to:

     A. Exponential increase in the amount of data available

     B. Increase in computing power and data storage capacity, at affordable cost

     C. All of the above

The correct answer is **C.**

The growth in big data and the machine learning revolution can be traced down to:

1. The availability of new datasets previously unavailable, such as sensor data from satellites, online activity of individuals, and the internet of things.
2. Advancement in computing power and data storage capacity (From kilobytes to petabytes)
3. Advancement in Machine Learning methods to analyze complex datasets, including programming languages such as Python, Java, SQL, etc.

A is incorrect. While the exponential increase in the amount of data available has indeed played a crucial role in the big data revolution, attributing the revolution solely to this factor overlooks the equally important contributions of advancements in computing power and data storage capacity. The availability of large datasets, such as sensor data from satellites, online activity logs, and the Internet of Things (IoT) devices, has certainly fueled the growth of big data. However, without the corresponding advancements in technology to process and store this data efficiently, the potential of these vast datasets could not have been fully realized.

B is incorrect. Similarly, attributing the big data revolution solely to the increase in computing power and data storage capacity, at affordable costs, does not provide a complete picture. While these technological advancements have been pivotal in enabling the processing and analysis of large datasets, the revolution also owes much to the exponential growth in the volume of data generated. This includes data from diverse sources such as social media, e-commerce, digital communications, and IoT devices. The synergy between the increased data availability and the technological capabilities to handle such data has been fundamental to the big data revolution.

CFA Level I, Quantitative Methods, Learning Module 11: Introduction to Big Data Techniques. LOS (c): Describe applications of Big Data and Data Science to investment management.

Q.3707 Sensor data is *most likely* taken from:

    A. individuals through their online activity such as product reviews, credit card purchases, and social media posts.

    B. businesses and corporations, including sales information, credit card data, and corporate exhaust.

    C. devices such as smartphones, cameras, and satellites.

The correct answer is **C.**

Sensor data is the output of a device that detects and responds to some type of input from the physical environment. Sensor data are collected from such devices as smartphones, cameras, RFID chips, and satellites. These devices are connected to computers via wireless networks.

A is incorrect. This option suggests that sensor data is taken from individuals through their online activities such as product reviews, credit card purchases, and social media posts. However, this type of data is more accurately described as user-generated content or digital footprint data, rather than sensor data. While this information can be valuable for various analyses, it does not originate from physical sensors or devices designed to detect changes in the environment. Instead, it is generated by individuals engaging with digital platforms and services, reflecting their preferences, opinions, and behaviors.

B is incorrect. This option indicates that sensor data comes from businesses and corporations, including sales information, credit card data, and corporate exhaust. Similar to option A, this type of data is not sensor data. Instead, it is transactional or operational data generated by business activities. While businesses may use sensor data as part of their operations (for example, in inventory management with RFID chips), the data described in this option pertains to business transactions and internal processes, not the output of devices sensing the physical environment.

CFA Level I, Quantitative Methods, Learning Module 11: Introduction to Big Data Techniques. LOS (a): Describe aspects of "fintech" that are directly relevant for the gathering and analyzing of financial data.

Q.3709 Machine learning is:

A. The autonomous acquisition of knowledge through the use of computer programs

B. The ability of machines to execute coded instructions

C. The selective acquisition of knowledge through the use of computer programs

The correct answer is **A.**

Machine learning refers to the autonomous acquisition of knowledge through computer programs such that the computers learn to work out solutions to problems without human intervention. Machine learning is the idea that computers have the ability to "learn" and execute changes independently.

B is incorrect. This option incorrectly defines machine learning as merely the ability of machines to execute coded instructions. While executing coded instructions is a fundamental capability of computers and machines, it does not encompass the concept of learning or the ability to improve performance on tasks over time without being explicitly programmed to do so. Traditional computer programs operate based on predefined rules and instructions provided by humans, without the ability to learn or adapt. In contrast, machine learning involves the development of algorithms that allow machines to learn from data and make decisions or predictions, thereby going beyond mere execution of coded instructions.

C is incorrect. This option describes machine learning as the selective acquisition of knowledge through the use of computer programs. While this description touches on an aspect of machine learning, it fails to capture the essence of autonomy in the learning process. Machine learning is not just about selectively acquiring knowledge; it is about doing so autonomously, without direct human intervention. The term "selective" might imply a more passive or limited approach to knowledge acquisition, whereas machine learning involves active analysis, interpretation, and learning from vast amounts of data in a way that improves the program's performance on specific tasks over time.

CFA Level I, Quantitative Methods, Learning Module 11: Introduction to Big Data Techniques. LOS (a): Describe aspects of "fintech" that are directly relevant for the gathering and analyzing of financial data.

Q.3710 Which of the following is most likely behind the increased adoption of automatic algorithmic trading?

     A. Increased efficiency

     B. Increased market destinations

     C. Ability to execute large trades

The correct answer is **B.**

Over time, financial markets have disintegrated into smaller markets consisting of electronic exchanges, alternative trading systems, and dark pools. Digital algorithms have made it possible to execute multiple trades over several global financial markets automatically. This has been their biggest selling point.

A is incorrect. While increased efficiency is a notable benefit of automatic algorithmic trading, it is not the primary reason behind its increased adoption. Efficiency in this context refers to the speed and accuracy with which trades are executed, minimizing slippage and operational costs. Although these are important factors, the core driver for the adoption of algorithmic trading is its ability to navigate and exploit opportunities across multiple market destinations, rather than efficiency alone.

C is incorrect. The ability to execute large trades is indeed a feature of algorithmic trading, particularly through techniques such as order slicing, which breaks down large orders into smaller, less market-disruptive transactions. However, this capability, while valuable, is not the primary reason for the increased adoption of algorithmic trading. The fragmentation of financial markets and the subsequent need to efficiently access and trade across these diverse venues are more significant factors driving the adoption of algorithmic trading strategies. The capacity to handle large trades, although beneficial, supports the broader objective of optimizing trade execution across multiple market destinations.

CFA Level I, Quantitative Methods, Learning Module 11: Introduction to Big Data Techniques. LOS (c): Describe applications of Big Data and Data Science to investment management.

Q.3712 A correct description of artificial intelligence is that:

A. It encompasses more advanced systems that are able to analyze information and make decisions based on machine-learning logic.

B. Its goals are very different from those of machine learning.

C. It terminates the need for human input in investment analysis.

The correct answer is **A.**

Fintech encompasses more advanced systems that are able to analyze information and make decisions based on machine-learning logic. Machines that "learn" how to perform tasks over time have been developed. The use of such systems has brought about high levels of inefficiency that surpass human capabilities.

B is incorrect. Fintech encompasses both Artificial intelligence (AI) and machine learning (ML). The two have largely similar goals: to create new and innovative products and services.

C is incorrect. It's widely accepted that AI and ML techniques are not self-sufficient. Human input is still an integral part of investment analysis.

CFA Level I, Quantitative Methods, Learning Module 11: Introduction to Big Data Techniques. LOS (c): Describe applications of Big Data and Data Science to investment management.