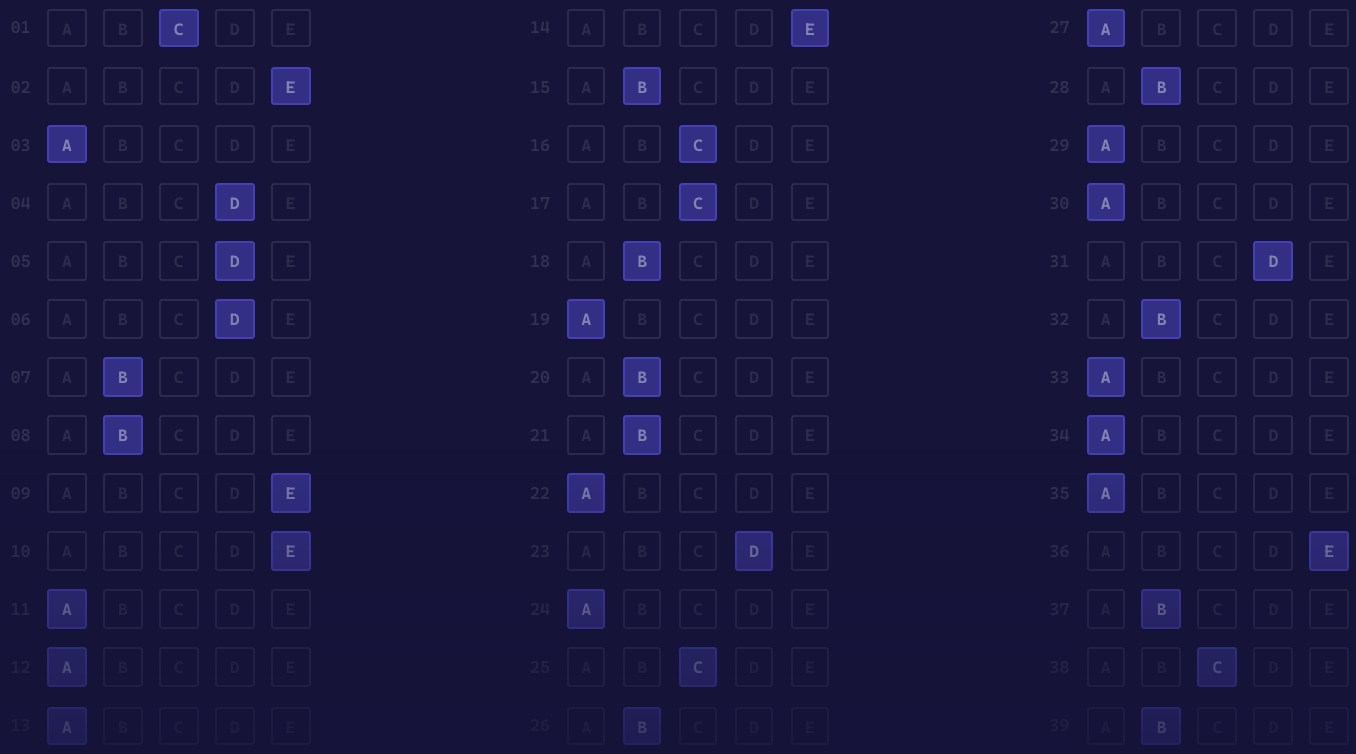
MOMETRIX
SMARTER PREP. BETTER RESULTS.



PRACTICE QUESTIONS

QuantInsti Quantitative Learning Private Limited

EPATN

EPAT News Sentiment and Alternative Data

EDITION	QUESTIONS	FORMAT
Demo	9	Limited Edition PDF

WEB www.mometrix.net

SUPPORT support@mometrix.net

Before you begin

What this is

9 practice questions for QuantInsti Quantitative Learning Private Limited `EPATN`, each with its answer key and an explanation. Answer a question before you read its key — the explanation is worth more once you have committed.

If something looks wrong

Send us three things and we will check it:

- the exam code, `EPATN`
- the question number
- the email address on your order

Write to **support@mometrix.net**. We reply within one business day.

mometrix.net

QUESTION 1 SINGLE CHOICE

In quantitative finance, which of the following best describes 'alternative data' as used in algorithmic trading?

A Data sourced exclusively from a stock exchange's official trade feed

B **Non-traditional datasets such as satellite imagery, social media, credit card transactions, and news used to gain an informational edge**

C Historical price and volume data only

D Audited financial statements filed by listed companies

ANSWER B

WHY Alternative data refers to non-traditional datasets beyond prices, volumes, and standard fundamentals. Common sources include satellite imagery, geolocation data, social media posts, credit-card transaction feeds, and news articles — exactly the set listed in option B. The other options describe traditional market data or fundamental data, which are not classified as alternative data.

QUESTION 2 SINGLE CHOICE

In news-based sentiment analysis, what does the term 'polarity' typically refer to?

A The grammatical part of speech of a word

B **The directional valence (positive, negative, neutral) of sentiment expressed in a text**

C The number of times a stock ticker is mentioned in an article

D The publication source credibility score

ANSWER B

WHY Polarity in sentiment analysis refers to the directional valence of expressed sentiment — whether the text is positive, negative, or neutral toward its subject (e.g., a company or stock). Mention counts, source credibility, and parts of speech are separate NLP features and are not what 'polarity' means.

mometrix.net

QUESTION 3 MULTIPLE CHOICE

Which of the following are common text preprocessing steps applied to raw news articles before sentiment scoring or feature extraction? (Choose two.)

- **A** **Tokenisation of the article text into individual words or sub-words**
- **B** **Removing stopwords such as 'the', 'is', and 'of'**
- **C** Converting every word to its uppercase form for consistency
- **D** Appending a unique timestamp to each news headline

ANSWER **A · B**

WHY Standard NLP preprocessing pipelines for news text typically include tokenisation (splitting text into words or sub-words) and stopword removal. Converting to uppercase is generally avoided because it destroys case-based cues such as proper nouns, and appending timestamps is a metadata operation, not a text-cleaning step.

QUESTION 4 SINGLE CHOICE

Which statement most accurately compares lexicon-based sentiment analysis with machine-learning-based sentiment classification for financial news?

- **A** Lexicon methods are always more accurate than ML methods on financial text
- **B** **Lexicon methods rely on predefined word–score dictionaries, while ML methods learn sentiment patterns from labelled training data**
- **C** ML methods cannot use lexicons at all
- **D** Lexicon methods require a GPU to run, whereas ML methods do not

ANSWER **B**

WHY Lexicon-based approaches (e.g., Loughran–McDonald, Vader-Finance) assign sentiment based on predefined dictionaries mapping words to polarity scores. Machine-learning approaches (e.g., logistic regression, LSTMs, transformers) learn sentiment boundaries from labelled corpora. Neither approach is universally more accurate, and ML methods can in fact incorporate lexicon-derived features.

mometrix.net

QUESTION 5 SINGLE CHOICE

The Loughran–McDonald sentiment word list is widely used for financial text. What is its primary intended purpose?

A Classifying sentiment of social-media chat messages

B **Providing finance-specific positive/negative word lists for analysing corporate filings, earnings releases, and financial news**

C Translating English financial terms into multiple languages

D Detecting insider-trading patterns in order-book data

ANSWER **B**

WHY Loughran and McDonald constructed finance-domain lexicons (Negative, Positive, Uncertainty, Litigious, StrongModal, WeakModal, Constraining) specifically to capture sentiment and tone in 10-K filings, earnings calls, and other corporate disclosures — not social chat, translation, or trade surveillance.

QUESTION 6 MULTIPLE CHOICE

Which of the following are valid examples of alternative data commonly used in quantitative trading? (Choose two.)

A Daily closing prices from a primary exchange

B **Satellite imagery of retail parking lots to estimate store traffic**

C **Aggregate anonymised credit-card transaction data**

D A company's audited annual report filed with the regulator

ANSWER **B · C**

WHY Satellite imagery (e.g., car counts at retail locations) and aggregated credit-card transaction panels are classic examples of alternative data because they are non-traditional, exogenous signals harvested from non-exchange sources. Daily closing prices and audited annual reports are traditional market and fundamental data, not alternative data.

mometrix.net

QUESTION 7 SINGLE CHOICE

In an event-study framework applied to news announcements, what does the 'abnormal return' represent?

A The total return earned by the stock on the event day

B The actual stock return minus the return expected under the normal performance model (e.g., market model)

C The risk-free rate of return over the event window

D The volume of shares traded on the event day

ANSWER B

WHY The abnormal return (AR) is defined as the difference between the realised return and the return predicted by a benchmark model such as the market model or CAPM. Cumulative abnormal returns (CARs) over a window are then tested for statistical significance to assess the market's reaction to the news event.

QUESTION 8 SINGLE CHOICE

Which of the following is a key challenge when using real-time news sentiment data in live trading models?

A All news headlines are written in exactly the same tone

B Sentiment scores are perfectly stable across time and sources

C Look-ahead bias, latency, and noisy/redundant signals can distort backtested and live performance

D Once a sentiment score is computed, it never needs updating

ANSWER C

WHY Live sentiment pipelines are vulnerable to look-ahead bias in backtests, varying latencies across news vendors, and high noise/redundancy (multiple headlines about the same event). Scores also drift as language, sources, and topics evolve, requiring periodic model refresh — none of which match the implausible statements in A, B, or D.

mometrix.net

QUESTION 9 SINGLE CHOICE

In natural language processing, what does TF-IDF (Term Frequency–Inverse Document Frequency) primarily capture?

A The grammatical role of each word in a sentence

B **A weighting that gives high scores to words that are frequent in a specific document but rare across the broader corpus**

C The exact semantic meaning of polysemous words

D The timestamp of when each word was first introduced into the language

ANSWER **B**

WHY TF-IDF combines term frequency (importance within one document) with inverse document frequency (rarity across the corpus), so words that are common everywhere (e.g., 'the') get down-weighted while words that distinguish a particular document get high weights. It does not encode grammar, exact semantics, or word-origin timestamps.

mometrix.net