Data Driven Methods in Finance: factors and datasets

Fall 2023: IEOR 4576

Naftali Cohen

Question/s of the week

Question: Compute the mean and variance of the random variable X = e^Z where Z is standard normal

Answer:

Introduction

Definition:

Factors are pivotal elements employed by quantitative managers to decipher stock returns. Essentially, they serve as variables that potentially illustrate the underlying reasons behind the fluctuations in stock returns. For instance, the hair color of a CEO could technically qualify as a factor, although it would ostensibly be an ineffective one. Conversely, the Price-to-Book (P/B) ratio of a company could also serve as a factor and is likely to be more pertinent and insightful.

Identifying Optimal Factors:

Selecting the most predictive factors is crucial as it empowers portfolio managers to construct more effective and profitable portfolios. Recognizing the right factors is instrumental in facilitating superior stock selection and consequently, optimizing portfolio performance.

Types of Factors in Quantitative Analysis:

- **Fundamental Factors:**

These encompass elements extracted from a firm's financial statements – the balance sheet, income statement, and cash flow statement. This category also includes any ratios formulated from these financial documents, providing insights into a company's financial health and performance.

- **Technical Factors:**

This type of factor involves analytical evaluations based on manipulations of a stock's price and volume data. Technical factors are crucial for analyzing market trends and predicting future price movements, typically without regard to financial statement analysis.

Types of Factors in Quantitative Analysis:

- **Economic Factors:**

Economic factors are macroeconomic elements influencing the overall economy, such as inflation rates, Gross Domestic Product (GDP), and unemployment rates. These factors play a significant role in assessing economic health and predicting market movements.

- **Alternative Factors:**

These factors are not encompassed within the previous categories and exhibit extensive diversity. Alternative factors may include unconventional and non-traditional data points that could have potential explanatory power for stock returns.

Understanding Fundamental Factors

Definition:
Fundamental Factors serve as indicators of a firm's intrinsic health, helping to assess its financial stability and performance.

Purpose:
They address pivotal questions such as:

- "Is the firm generating substantial profit?"
- "Is there a looming risk of the firm declaring bankruptcy?"

Source:
These factors are typically derived from meticulous analysis of a company's financial statements, often intertwined with relevant stock market data to provide a comprehensive overview of the firm's financial condition.

Categories:
Fundamental Factors can be meticulously organized into seven distinctive subcategories:

Fundamental Factors: Valuation Factors

Valuation factors, such as the P/E and P/B ratios, are pivotal in assessing whether a stock is priced fairly, appearing relatively cheap or expensive, and they fundamentally reflect the perceived value of a company in the market. They provide a lens through which investors can make informed decisions about the attractiveness of a stock, given its current price relative to intrinsic measures of value.

Ticker	Dividend Yield (DY)	Price-to- Book Ratio (P/B)	Price-to-Cash-Flow Ratio (P/CF)	Price-to-Earnings Ratio (P/E)	Price-to-Earnings-to-Growth Ratio (PEG)	Price-to-Earnings-to-Growth-plus-Yield Ratio (PEGY)	Price-to-Sales Ratio (P/S)
AAPL	0.61	34.53	27.96	39.30	3.24	3.08	8.22
MSFT	0.92	13.63	25.40	35.41	2.45	2.30	11.43
AMZN	0.00	19.74	29.56	94.04	2.59	2.59	4.70
GOOGL	0.00	5.19	19.43	30.94	1.87	1.87	6.43
TSLA	0.00	41.73	153.81	1,203.07	2.95	2.95	23.74
FB	0.00	5.58	19.43	25.98	1.57	1.57	8.31
BRK.B	0.00	1.31	13.16	15.17	N/A	N/A	1.95
JNJ	2.47	6.43	19.19	24.39	5.66	3.60	5.12
WMT	1.48	5.42	12.35	22.79	3.38	2.77	0.75
JPM	2.86	1.61	12.40	15.18	N/A	N/A	2.91

*Data for the end of 2020

Fundamental Factors: Valuation Factors

1. Dividend Yield:

- **Definition**: A financial ratio that shows how much a company pays out in dividends each year relative to its stock price.
- **Formula**: Annual Dividends per Share / Price per Share
- **Use**: It is often used by investors looking for companies with a steady income through dividends, typically reflecting stability and profitability.

2. P/B (Price-to-Book Ratio):

- **Definition**: A financial ratio used to compare a company's current market value to its book value (Book Value = Total assets - intangible assets - total liabilities).
- **Formula**: Market Price per Share / Book Value per Share
- **Use**: It helps in identifying undervalued stocks and assessing the inherent value of a firm.

Fundamental Factors: Valuation Factors

3. P/CF (Price-to-Cash Flow Ratio):

- **Definition**: A ratio used to compare a company's market value to its cash flow.
- **Formula**: Stock Price / Cash Flow per Share
- **Use**: It is often used to assess a company's financial health, with a lower ratio suggesting better value.

4. P/E (Price-to-Earnings Ratio):

- **Definition**: A valuation ratio of a company's current share price compared to its per-share earnings.
- **Formula**: Market Value per Share / Earnings per Share (EPS)
- **Use**: Commonly used to assess the relative valuation of companies and growth expectations.

Fundamental Factors: Valuation Factors

5. PEG (Price/Earnings to Growth Ratio):

- **Definition**: A stock's P/E ratio divided by the growth rate of its earnings.
- **Formula**: P/E Ratio / Earnings Growth Rate
- **Use**: It is used to determine a stock's value while considering the company's earnings growth, with a lower ratio indicating a potentially undervalued stock.

6. PEGY (Price/Earnings to Growth and Dividend Yield):

- **Definition**: Similar to PEG but also accounts for dividend yield.
- **Formula**: P/E Ratio / (Earnings Growth Rate + Dividend Yield)
- **Use**: It gives a more comprehensive picture of a stock's value, growth, and yield, providing insights into dividend-paying stocks.

Fundamental Factors: Valuation Factors

7. P/S (Price-to-Sales Ratio):

- **Definition**: A valuation ratio that compares a company's stock price to its revenues.
- **Formula**: Market Capitalization / Total Sales or Revenues
- **Use**: It is useful for assessing the value of companies with unstable earnings, providing insights into the revenue-generating ability relative to the stock price.

Fundamental Factors: Size Factors

Size factors, including common equity, market capitalization, and total assets, represent the scale and extent of a company's operations and assets. They are crucial for investors to understand the company's position in its industry and its capability to leverage scale, which often correlates with stability and resilience in various market conditions.

Ticker	Common Equity (CE)	Market Capitalization (SIZE)	Total Assets (TA)
AAPL	65.34	2,255.97	323.89
MSFT	123.39	1,681.61	301.00
AMZN	82.77	1,634.17	282.18
GOOGL	212.92	1,104.81	299.24
TSLA	16.03	668.91	45.69
FB	117.73	656.67	146.44
BRK.B	415.15	543.61	829.95
JNJ	64.47	414.31	170.69
WMT	75.31	407.84	237.38
JPM	241.05	387.34	3,246.07

*Data for the end of 2020

Fundamental Factors: Size Factors

1. Common Equity:

- **Definition**: Common Equity represents the net worth of a company, indicating the owners' residual claim on assets in the event of liquidation.
- **Formula**: Common Equity = Total Assets - Total Liabilities
- **Use**: Common Equity is crucial for evaluating a company's financial stability and risk, giving investors insights into the company's ability to withstand economic downturns and potential for return on investment.

2. Market Capitalization:

- **Definition**: Market Capitalization is a measurement of the size and value of a company, representing the total market value of its outstanding shares of stock.
- **Formula**: Market Capitalization = Stock Price × Number of Outstanding Shares
- **Use**: It helps in investment decisions, categorizing stocks into different sizes (small-cap, mid-cap, and large-cap) and determining the company's risk and return profile, stability, and growth potential.

Fundamental Factors: Size Factors

3. Total Assets:

- **Definition**: Total Assets are the sum of all tangible and intangible assets owned by a company, representing the total resources available for generating revenues.
- **Formula**: Total Assets = Total Liabilities + Total Equity
- **Use**: Analyzing Total Assets is central for assessing a company's size, financial health, and operational efficiency, helping investors evaluate how effectively a company is utilizing its assets to generate profits.

Fundamental Factors: Operational Efficiency Factors

Operational efficiency factors, like equity and asset turnovers, are metrics that illustrate how effectively a company utilizes its assets to generate sales and profits. These factors are integral for investors aiming to evaluate a company's management effectiveness and its ability to optimize resource use for maximum productivity and profitability.

Ticker	Equity Turnover (ET)	Fixed Asset Turnover (FAT)	Inventory Turnover (IT)	Total Asset Turnover (TAT)
AAPL	3.52	6.64	38.82	0.85
MSFT	1.28	2.85	12.84	0.49
AMZN	5.00	3.09	8.67	1.23
GOOGL	0.84	1.98	59.42	0.57
TSLA	2.37	1.35	5.07	0.62
FB	0.75	1.71	N/A	0.54
BRK.B	0.68	1.58	10.55	0.34
JNJ	1.32	4.51	2.18	0.47
WMT	6.85	4.36	9.32	2.28
JPM	0.50	5.14	4.09	0.04

*Data for the end of 2020

Fundamental Factors: Operational Efficiency Factors

1. Equity Turnover:

- **Definition**: Equity Turnover measures how efficiently a company uses equity to generate sales and is a critical indicator of operating performance.
- **Formula**: Equity Turnover = Sales / Average Shareholders' Equity
- **Use**: It aids investors and analysts in assessing how effectively a company is utilizing its equity to generate revenue, with higher ratios indicating greater efficiency and profitability.

2. Fixed Asset Turnover:

- **Definition**: Fixed Asset Turnover quantifies a company's ability to generate sales from its fixed assets such as property, plant, and equipment.
- **Formula**: Fixed Asset Turnover = Sales / Average Net Fixed Assets
- **Use**: It is instrumental in evaluating how efficiently a company is using its fixed assets to generate sales. A higher ratio indicates better utilization of fixed assets in revenue generation.

Fundamental Factors: Operational Efficiency Factors

3. Inventory Turnover:

- **Definition**: Inventory Turnover gauges how many times a company's inventory is sold and replaced over a specific period, signaling the efficiency of inventory management.
- **Formula**: Inventory Turnover = Cost of Goods Sold / Average Inventory
- **Use**: This ratio is pivotal for companies to assess how effectively they are managing their inventory. High inventory turnover usually signifies strong sales and efficient inventory management.

4. Total Asset Turnover:

- **Definition**: Total Asset Turnover is a measure of a company's ability to use its assets to generate sales, indicative of operational efficiency.
- **Formula**: Total Asset Turnover = Sales / Average Total Assets
- **Use**: It helps in assessing how effectively a company is utilizing all of its assets to generate revenue. A higher ratio implies better asset utilization and operational efficiency.

Fundamental Factors: Solvency Factors

Solvency factors, including ratios like the current and quick ratios, evaluate a company's long-term financial sustainability and its ability to meet its long-term obligations. They are essential for risk assessment, enabling investors to gauge the financial health and stability of a company, thus avoiding those with a high likelihood of bankruptcy.

Ticker	Cash-Flow-from-Operations Ratio (CFOR)	Cash Ratio (CR)	Current Ratio (CUR)	Quick Ratio (QR)
AAPL	0.77	0.86	1.36	1.22
MSFT	0.94	1.97	2.53	2.30
AMZN	0.54	0.67	1.11	0.84
GOOGL	1.18	2.75	3.41	3.28
TSLA	0.33	1.11	1.63	1.24
FB	2.83	4.68	5.51	5.35
BRK.B	N/A	N/A	N/A	N/A
JNJ	0.56	0.79	1.48	1.17
WMT	0.40	0.21	0.79	0.27
JPM	N/A	N/A	N/A	N/A

*Data for the end of 2020

Fundamental Factors: Solvency Factors

1. Cash Flow from Operations Ratio:

- **Definition**: This ratio assesses a company's ability to generate cash from its operating activities, relative to its liabilities.
- **Formula**: Cash Flow from Operations Ratio = Cash Flows from Operations / Current Liabilities
- **Use**: It is instrumental in evaluating a company's liquidity and its ability to meet its short-term obligations. A higher ratio indicates a better ability to cover current liabilities with cash generated from operations.

2. Cash Ratio:

- **Definition**: The Cash Ratio measures a company's ability to pay off its current liabilities using only cash and cash equivalents.
- **Formula**: Cash Ratio = (Cash + Cash Equivalents) / Current Liabilities
- **Use**: This is a stringent liquidity measure that indicates the extent to which a company can immediately settle its current liabilities. A higher ratio means higher liquidity, reflecting a better immediate paying capacity.

Fundamental Factors: Solvency Factors

3. Current Ratio:

- **Definition**: The Current Ratio is a liquidity ratio that measures whether a firm has enough resources to meet its short-term obligations.
- **Formula**: Current Ratio = Current Assets / Current Liabilities
- **Use**: It is widely used to gauge a company's financial health and liquidity. A ratio above 1 implies that the company can pay off its short-term obligations if they came due at that point.

4. Quick Ratio:

- **Definition**: Also known as the acid-test ratio, the Quick Ratio assesses a company's ability to meet its short-term obligations using its most liquid assets.
- **Formula**: Quick Ratio = (Current Assets - Inventories) / Current Liabilities
- **Use**: This ratio provides a more stringent view of a company's short-term liquidity than the current ratio, as it excludes inventories from assets. A quick ratio of 1 or more is usually considered healthy, signifying that the company can fulfill its immediate obligations without relying on the sale of inventory.

Fundamental Factors: Financial Risk Factors

Financial risk factors, such as the debt to equity and interest coverage ratios, provide insights into the financial risks a company is undertaking, primarily related to its leverage and debt management. They enable investors to understand the risk levels associated with the company's financial structure and strategies, assessing the balance between risk and reward.

Ticker	Cash-Flow Coverage Ratio (CFCR)	Debt-to-Equity Ratio (D/E)	Interest Coverage Ratio (ICR)	Total Debt Ratio (TDR)
AAPL	0.66	3.96	23.35	0.38
MSFT	0.80	1.44	22.21	0.28
AMZN	0.58	2.41	12.12	0.34
GOOGL	2.06	0.41	127.29	0.09
TSLA	0.29	1.67	1.33	0.33
FB	3.03	0.24	N/A	0.08
BRK.B	0.38	0.98	11.72	0.13
JNJ	0.57	1.65	95.82	0.22
WMT	0.48	1.92	9.50	0.29
JPM	0.06	10.97	N/A	0.17

*Data for the end of 2020

Fundamental Factors: Financial Risk Factors

1. Cash Flow from Operations Ratio:

- **Definition**: This ratio assesses a company's ability to generate cash from its operating activities, relative to its liabilities.
- **Formula**: Cash Flow from Operations Ratio = Cash Flows from Operations / Current Liabilities
- **Use**: It is instrumental in evaluating a company's liquidity and its ability to meet its short-term obligations. A higher ratio indicates a better ability to cover current liabilities with cash generated from operations.

2. Cash Ratio:

- **Definition**: The Cash Ratio measures a company's ability to pay off its current liabilities using only cash and cash equivalents.
- **Formula**: Cash Ratio = (Cash + Cash Equivalents) / Current Liabilities
- **Use**: This is a stringent liquidity measure that indicates the extent to which a company can immediately settle its current liabilities. A higher ratio means higher liquidity, reflecting a better immediate paying capacity.

Fundamental Factors: Financial Risk Factors

3. Current Ratio:

- **Definition**: The Current Ratio is a liquidity ratio that measures whether a firm has enough resources to meet its short-term obligations.
- **Formula**: Current Ratio = Current Assets / Current Liabilities
- **Use**: It is widely used to gauge a company's financial health and liquidity. A ratio above 1 implies that the company can pay off its short-term obligations if they came due at that point.

4. Quick Ratio:

- **Definition**: Also known as the acid-test ratio, the Quick Ratio assesses a company's ability to meet its short-term obligations using its most liquid assets.
- **Formula**: Quick Ratio = (Current Assets - Inventories) / Current Liabilities
- **Use**: This ratio provides a more stringent view of a company's short-term liquidity than the current ratio, as it excludes inventories from assets. A quick ratio of 1 or more is usually considered healthy, signifying that the company can fulfill its immediate obligations without relying on the sale of inventory.

Fundamental Factors: Corporate Activity Factors

Corporate activity factors encompass elements like stock buybacks and R&D expenditure, reflecting a company's strategic initiatives and internal confidence in its future prospects. These factors offer crucial insights into the company’s strategic focus and investment in innovation and growth, signaling the management's view of future opportunities and challenges.

Fundamental Factors: Corporate Activity Factors

1. Stock Buybacks:

- **Definition**: This refers to the repurchasing of shares by the company that issued them, a way companies return value to shareholders.
- **Formula**: No direct formula, it's quantified as the number of shares bought multiplied by the purchase price.
- **Use**: It is used to analyze a company's investment in itself. It can indicate the company's financial health and its belief in its own value, often leading to a rise in stock price due to reduced share count, potentially increasing earnings per share (EPS).

2. Insider Purchases:

- **Definition**: This refers to the buying of the company's stock by individuals who work for the company, often executives or directors.
- **Formula**: It is quantified by the number of shares bought by insiders, but no standard formula exists.
- **Use**: It is used to gauge the confidence of the company's insiders in the firm's future prospects. A high volume of insider purchases may signal to external investors that the insiders believe the company will perform well in the future.

Fundamental Factors: Corporate Activity Factors

3. Research and Development (R&D) Expenditure:

- **Definition**: This represents the money spent by a company on activities to innovate and introduce new products or services.
- **Formula**: R&D Expenditure is a line item on the income statement, and no direct formula is used to calculate it.
- **Use**: It is critical for evaluating a company's future growth potential, especially in technology and pharmaceutical sectors. High R&D spending may imply a strong commitment to innovation and the development of new products, services, or technologies, potentially leading to long-term competitive advantages and higher future revenues.

Technical Factors

Technical factors, derived from historical price and volume data, encompass a variety of indicators such as opening, high, low, and closing prices, volume, open interest, and bid and ask prices. These factors are pivotal for tracking short-term fluctuations and discerning patterns in stocks' relative value due to their frequent updating, enabling more real-time insight compared to fundamental factors.

Technical factors can be broadly categorized into four subcategories, namely:

- **Liquidity Risk Factors**: Evaluating the ease of executing trades without affecting the asset's price.
- **Price-based Factors**: Analysing price patterns and trends to forecast future movements.
- **Volume-based Factors**: Examining trading volumes to understand market participants' behavior and market strength.
- **Overall Market Movement Factors**: Assessing the broader market trends and movements to gauge the market sentiment.

Technical Factors: Liquidity Risk Factors

1. **Amihud Illiquidity:**
   - **Definition**: Measures the price impact of order flow, representing the daily price response associated with one dollar of trading volume (the number of shares traded during a trading day).
   - **Formula**: Amihud Illiquidity = |Return|/Volume
   - **Use**: Utilized to identify assets that may have higher transaction costs and can be employed to adjust trading strategies to mitigate the impact of trades on asset prices.

2. Invariance Illiquidity:
   - **Definition**: This factor gauges the resilience of the market or how quickly the price impact of trades reverts following a liquidity shock.
   - **Formula**: It is often quantified using advanced econometric models, measuring the speed and extent of price reversion after a trade.
   - **Use**: Employed to assess the stability and recovery of markets post-liquidity events, aiding in the formulation of strategies to capitalize on price recovery or minimize adverse impacts.

Technical Factors: Liquidity Risk Factors

3. Trading Turnover:

- **Definition**: It represents the ratio of average trading volume to market capitalization, reflecting the asset's trading activity.
- **Formula**: Trading Turnover= Market Capitalization/ Average Trading Volume
- **Use**: Trading turnover is instrumental in identifying highly traded assets, which may have lower transaction costs, and in discerning assets' liquidity profiles, which helps in optimizing trade execution.

Technical Factors: Price-Based Factors

1. Beta:

- **Definition**: Beta measures the sensitivity of an asset's returns to the returns of the market or a benchmark index.
- **Formula**: Beta = Cov(R_e, R_m)/Var(R_m) where R_e is the asset return and R_m is the market return.
- **Use**: Beta is crucial for assessing market risk of an asset, guiding asset allocation decisions, and informing about the asset's contribution to portfolio risk.

2. Bollinger Bands:

- Definition: Bollinger Bands consist of a middle band being an N-period simple moving average (SMA), and an upper and lower band at k times an N-period standard deviation above and below the SMA.
- **Formula**:
  - Middle Band: SMA(N)
  - Upper Band: SMA(N) + (k * σ)
  - Lower Band: SMA(N) - (k * σ)
  - σ is the standard deviation of the price over N periods.
- **Use**: These bands aid in identifying overbought or oversold conditions in the market, helping traders to spot potential reversals and trend continuations.

Technical Factors: Price-Based Factors

3. Twelve-Month Momentum:

- **Definition**: It is the rate of acceleration of an asset's price or volume, typically calculated by comparing the current closing price to the closing price N periods ago.
- **Formula**: Momentum = Current Closing Price - N periods ago Closing Price
- **Use**: Used for identifying the strength of an asset's trend, enabling traders to capitalize on sustained movements in asset prices.

4. Relative Strength Index (RSI)

- **Definition**: RSI is a momentum oscillator that measures the speed and change of price movements. It is typically used to identify overbought or oversold conditions.
- **Formula**: RSI = 100 - 100/(1 + RS) where RS = Average Gain/ Average Loss over a specified period
- **Use**: RSI helps traders identify potential reversal points in the market, acting as a cue to initiate or close positions based on perceived market strength or weakness.

Technical Factors: Volume-Based Factors

1. Short Interest Ratio

- **Definition**: The Short Interest Ratio represents the ratio of the total number of shares sold short to the average daily trading volume, highlighting the level of investor pessimism or bearishness on a stock.
- **Formula**: Short Interest Ratio= Average Daily Trading Volume/ Total Number of Shares Shorted
- **Use**: It helps in assessing market sentiment and potential price movements, as a high ratio can indicate that a price increase may be forthcoming due to short-sellers covering their positions, leading to buying pressure.

2. Short Interest Ratio Minus One Year Average

- **Definition**: This factor represents the deviation of the current short interest ratio from its one-year average, serving as an indicator of recent shifts in investor sentiment relative to longer-term behavior.
- **Formula**: Short Interest Ratio − One Year Average Short Interest Ratio
- **Use**: This relative measure assists in detecting significant changes in market sentiment over the short term, providing insights into potential price movements and investment opportunities based on comparative investor pessimism or optimism.

Technical Factors: Market Movement Factors

1. Market Beta

- **Definition**: Market Beta is a measure of a stock's volatility in relation to the overall market. It is used to assess the systematic risk of a security or a portfolio in relation to the market as a whole.
- **Formula**: β = Covariance(StockReturn,MarketReturn)/ Variance(MarketReturn)
- **Use**: A beta greater than 1 suggests that the stock is more volatile than the market, while a beta less than 1 indicates less volatility. It helps investors in constructing diversified portfolios and managing risk, allowing the adjustment of exposure to market movements.

2. Moving Average Convergence Divergence (MACD)

- **Definition**: MACD is a trend-following momentum indicator that shows the relationship between two moving averages of a security's price, typically used to identify potential buy or sell signals.
- **Formula**: MACD=12-days EMA - 26-days EMA
- **Use**: When the MACD line crosses above the signal line, it is a buy signal, and when it crosses below, it is a sell signal. It is used to spot changes in the strength, direction, momentum, and duration of a trend in a stock's price.

Economic Factors

Economic factors comprise variables that influence the macroeconomy and, subsequently, stock returns. These include renowned indicators like

- Gross Domestic Product (GDP) growth,
- the slope of the yield curve,
- unemployment rates, and
- inflation levels.

These variables often find a place in macroeconomic models of stock returns as they possess the potential to impact substantial segments of the market. Portfolio managers must exercise caution when incorporating macroeconomic data into their models, given the inherent time lags in data availability and reporting. A thoughtful approach requires the acknowledgment and adjustment for such delays to ensure the relevance and accuracy of the models developed.

Alternative Factors

Alternative factors encompass a unique set of elements in financial analysis and modeling that stand distinct from the conventional classifications of fundamental, technical, and economic determinants. Within this domain, three prominent subcategories emerge:

- **Analyst Factors**: Derived from expert predictions and assessments.
- **Captured Factors**: Specialized metrics, often proprietary in nature.
- **Social Responsibility Factors**: Indicators that gauge a company's commitment to ethical, sustainable, and community-centric practices.

Ticker	Median Rating (ARMR)	Percent Buys (ARPB)	Percent Sells (ARPS)	SD of Rating (ARSD)	Number of Analysts (EPSFN)
AAPL	2.00	75.00	5.00	0.90	29.00
MSFT	2.00	91.43	0.00	0.64	27.00
AMZN	2.00	94.00	0.00	0.58	40.00
GOOGL	2.00	92.31	0.00	0.65	29.00
TSLA	3.00	37.14	31.43	1.32	15.00
FB	2.00	88.24	3.92	0.76	40.00
BRK.B	2.50	50.00	0.00	0.96	3.00
JNJ	2.00	72.22	0.00	0.80	14.00
WMT	2.00	79.41	2.94	0.81	28.00
JPM	2.00	62.96	7.41	0.74	22.00

Alternative Factors: Analyst Factors

1. Median Rating:

- **Definition**: Represents the middle value of collected analyst recommendations, providing a measure to understand the consensus opinion about a stock.
- **Formula**: When analyst ratings are numerically expressed (1=Buy, 5=Sell), the median rating is found by ordering the ratings and finding the middle one. If there is an even number of ratings, the median is the average of the two middle numbers.
- **Use**: This is used by investors to gauge the overall analyst sentiment for a stock and to identify potential overvalued or undervalued stocks.

2. Percent Buys/Sells:

- **Definition**: Represents the proportion of analysts recommending to buy or sell a stock.
- **Formula**: (Number of Buy Recommendations / Total Number of Recommendations) * 100 for buys and similarly for sells.
- **Use**: This helps in understanding the consensus of market opinion on whether the stock should be bought or sold, informing investment decisions.

Alternative Factors: Analyst Factors

3. Standard Deviation of Rating:

- **Definition**: Measures the amount of variance or dispersion of a set of analyst ratings.
- **Formula**: The standard deviation of the rating...
- **Use**: It is used to assess the level of disagreement among analysts; a higher standard deviation indicates a wider range of opinions, which could imply higher uncertainty about the stock's future.

4. Number of Analysts:

- **Definition**: Refers to the total number of analysts providing recommendations, ratings, or coverage for a stock.
- **Formula**: Count of individual analysts providing a rating or recommendation for the stock.
- **Use**: Used to determine the breadth of opinion and consensus on a stock; more analysts might suggest broader market interest and potentially more reliable consensus, though it may also lead to more varied opinions.

Alternative Factors: Captured Factors

Captured factors encompass a spectrum of data acquired through various sources including GPS, satellites, social media, and news feeds, which are collected using sophisticated programs and technologies. These are available with higher frequency and are often accessed earlier than typical fundamental data, presenting a more immediate reflection of current circumstances.

Method of Acquisition:

Web scraping is a predominant method for acquiring such captured data, enabling the extraction of real-time and diverse information directly from the web pages of various domains, whether they are company sites, consumer sites, or social media platforms.

Alternative Factors: Captured Factors

1. Transaction & Consumer Spending Data:

Entities such as Sandalwood Advisors, Skupos, and 90 West offer insights into consumer spending patterns. Sandalwood focuses on Chinese consumer spending on major eCommerce platforms, Skupos provides data on consumer spending in convenience stores, and 90 West provides traditional transactional data and tracks US gift card spending.

2. Media Consumption & Advertising Data:

Insightful data on media consumption and advertising spending is provided by entities like Standard Media Index and Netflix. Standard Media Index offers detailed data on advertising spending across various media platforms, and Netflix discloses data on global viewing trends and preferences.

Alternative Factors: Captured Factors

3. App & Online Activity Data:
Platforms like Sensor Tower and Similar Web stand out for their focus on app usage and online traffic data. Sensor Tower specializes in app store data and trends, while Similar Web provides insights into website traffic patterns and clickstream-level data.

4. Healthcare & Pharmaceutical Data:
Entities such as MedMine and Kyber Data Science contribute data and insights related to the healthcare and pharmaceutical industries. MedMine leverages large panels to track sales metrics for medical device and pharma distributors, and Kyber Data Science extracts data from medical and pharmacy claims.

Alternative Factors: Captured Factors

5. Economic & Financial Market Data:

FRED and S3 Partners play pivotal roles in delivering economic and equity market data. FRED provides a wealth of macroeconomic indicators, while S3 Partners specializes in offering short interest data across global equity markets, with minimal lag.

6. Employment & Human Resource Data:

Revelio Labs is distinctive for providing extensive human resource data, covering thousands of global companies and offering insights into diverse dimensions like headcount, job postings, and employee surveys.

Alternative Factors: Captured Factors

7. Travel & Hospitality Data:
For the travel and hospitality sectors, OpenTable and TSA are notable. OpenTable reveals restaurant reservation volumes across its network, while TSA provides data on passenger volumes passing through US airport checkpoints.

8. Market Estimates & Analyst Data:
Visible Alpha excels in delivering detailed consensus estimates at the KPI level, syncing and standardizing underlying analyst models to offer a unified view into market expectations.

Alternative Factors: Captured Factors

9. Market Research & Consumer Behavior Data:
Companies like M Science and Pacific Epoch provide crucial market research and insights into consumer behaviors across different regions and industries, focusing on alternative data-based research and consumer demand trends, particularly in Asia.

10. Technology Data:
GWS Solutions is highlighted for its insight into network coverage quality, offering data and analytics in network performance and reliability.

Alternative Factors: Data Coverage by Sector

Alternative Factors: Data Coverage by Sector












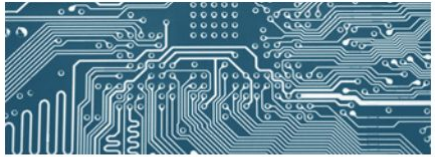



Alternative Factors: Data Coverage by Sector

1. Retail (sale of goods and services to consumers)**:**

Alternative Factors: Data Coverage by Sector

1. Retail (sale of goods and services to consumers)**:**
Datasets for this sector might encompass sales data, customer footfall, inventory levels, and consumer spending patterns, mainly focusing on brick-and-mortar stores and online retailers.

2. Restaurants:

Alternative Factors: Data Coverage by Sector

1. Retail (sale of goods and services to consumers)**:**
Datasets for this sector might encompass sales data, customer footfall, inventory levels, and consumer spending patterns, mainly focusing on brick-and-mortar stores and online retailers.

2. Restaurants:
In this sector, datasets may include reservation volumes, customer reviews, and spending patterns in different types of restaurants, ranging from fast-casual chains to high-end establishments.

3. Internet, eCommerce:

Alternative Factors: Data Coverage by Sector

1. Retail (sale of goods and services to consumers)**:**
Datasets for this sector might encompass sales data, customer footfall, inventory levels, and consumer spending patterns, mainly focusing on brick-and-mortar stores and online retailers.

2. Restaurants:
In this sector, datasets may include reservation volumes, customer reviews, and spending patterns in different types of restaurants, ranging from fast-casual chains to high-end establishments.

3. Internet, eCommerce:
Data here can include user behavior, traffic, and sales data from online stores and platforms, illuminating trends in online consumer behavior and preferences.

4. Media:

Alternative Factors: Data Coverage by Sector

1. Retail (sale of goods and services to consumers)**:**
Datasets for this sector might encompass sales data, customer footfall, inventory levels, and consumer spending patterns, mainly focusing on brick-and-mortar stores and online retailers.

2. Restaurants:
In this sector, datasets may include reservation volumes, customer reviews, and spending patterns in different types of restaurants, ranging from fast-casual chains to high-end establishments.

3. Internet, eCommerce:
Data here can include user behavior, traffic, and sales data from online stores and platforms, illuminating trends in online consumer behavior and preferences.

4. Media:
Datasets would capture insights into media consumption patterns, advertising spends, subscriber counts, content popularity, and audience demographics across various media platforms.

Alternative Factors: Data Coverage by Sector

5. Communications:

Alternative Factors: Data Coverage by Sector

5. Communications:

This sector's data can highlight network performance, subscriber data, usage patterns, and customer satisfaction levels related to telecom services.

6. Video Games:

Alternative Factors: Data Coverage by Sector

5. Communications:
This sector's data can highlight network performance, subscriber data, usage patterns, and customer satisfaction levels related to telecom services.

6. Video Games:
Datasets in this sector can provide information on game sales, user engagement levels, in-game purchases, and trends in player behavior.

7. Software:

Alternative Factors: Data Coverage by Sector

5. Communications:
This sector's data can highlight network performance, subscriber data, usage patterns, and customer satisfaction levels related to telecom services.

6. Video Games:
Datasets in this sector can provide information on game sales, user engagement levels, in-game purchases, and trends in player behavior.

7. Software:
Data can cover software sales, user adoption rates, subscription numbers, usage patterns, and customer feedback on various software products and services.

8. Semiconductors:

Alternative Factors: Data Coverage by Sector

5. Communications:
This sector's data can highlight network performance, subscriber data, usage patterns, and customer satisfaction levels related to telecom services.

6. Video Games:
Datasets in this sector can provide information on game sales, user engagement levels, in-game purchases, and trends in player behavior.

7. Software:
Data can cover software sales, user adoption rates, subscription numbers, usage patterns, and customer feedback on various software products and services.

8. Semiconductors:
Datasets can offer insights into production volumes, sales data, supply chain information, and market demand for semiconductor products.

Alternative Factors: Data Coverage by Sector

9. Construction, Housing, Real Estate:

Alternative Factors: Data Coverage by Sector

9. Construction, Housing, Real Estate:
Data might cover construction activity levels, housing sales and prices, real estate listings and transactions, and market trends in real estate.

10. Airlines:

Alternative Factors: Data Coverage by Sector

9. Construction, Housing, Real Estate:
Data might cover construction activity levels, housing sales and prices, real estate listings and transactions, and market trends in real estate.

10. Airlines:
Data can include passenger volumes, flight data, revenue from ticket sales, customer satisfaction, and other operational data related to airlines.

11. Automotive:

Alternative Factors: Data Coverage by Sector

9. Construction, Housing, Real Estate:
Data might cover construction activity levels, housing sales and prices, real estate listings and transactions, and market trends in real estate.

10. Airlines:
Data can include passenger volumes, flight data, revenue from ticket sales, customer satisfaction, and other operational data related to airlines.

11. Automotive:
Datasets might encompass sales volumes, production levels, inventory data, and consumer preferences in the automotive industry.

12. Chemicals:

Alternative Factors: Data Coverage by Sector

9. Construction, Housing, Real Estate:
Data might cover construction activity levels, housing sales and prices, real estate listings and transactions, and market trends in real estate.

10. Airlines:
Data can include passenger volumes, flight data, revenue from ticket sales, customer satisfaction, and other operational data related to airlines.

11. Automotive:
Datasets might encompass sales volumes, production levels, inventory data, and consumer preferences in the automotive industry.

12. Chemicals:
Data can highlight production volumes, sales data, market demand, and supply chain information for chemicals and related products.

Alternative Factors: Data Coverage by Sector

13. Industrials:

Alternative Factors: Data Coverage by Sector

13. Industrials:
In this sector, datasets may include production levels, operational data, sales, and supply chain information for industrial goods and manufacturing entities.

14. Healthcare Devices & Pharmaceutical:

Alternative Factors: Data Coverage by Sector

13. Industrials:
In this sector, datasets may include production levels, operational data, sales, and supply chain information for industrial goods and manufacturing entities.

14. Healthcare Devices & Pharmaceutical:
Datasets can cover sales of medical devices and pharmaceuticals, clinical trial data, medical claims data, and regulatory information.

15. Hotels, Tourism:

Alternative Factors: Data Coverage by Sector

13. Industrials:
In this sector, datasets may include production levels, operational data, sales, and supply chain information for industrial goods and manufacturing entities.

14. Healthcare Devices & Pharmaceutical:
Datasets can cover sales of medical devices and pharmaceuticals, clinical trial data, medical claims data, and regulatory information.

15. Hotels, Tourism:
Data in this sector can include booking volumes, occupancy rates, customer reviews, and spending patterns in hotels and other tourism-related services.

16. Macroeconomic:

Alternative Factors: Data Coverage by Sector

13. Industrials:
In this sector, datasets may include production levels, operational data, sales, and supply chain information for industrial goods and manufacturing entities.

14. Healthcare Devices & Pharmaceutical:
Datasets can cover sales of medical devices and pharmaceuticals, clinical trial data, medical claims data, and regulatory information.

15. Hotels, Tourism:
Data in this sector can include booking volumes, occupancy rates, customer reviews, and spending patterns in hotels and other tourism-related services.

16. Macroeconomic:
Datasets in this sector provide insights into economic indicators such as GDP growth, unemployment rates, inflation rates, and other macro-level economic trends and statistics.

Alternative Factors: Data Coverage by Sector

17. Financial Market Data:
This sector's datasets offer insights into stock prices, trading volumes, market capitalization, and other relevant financial market information.

18. Consumer Staples (essential products used by consumers)**:**
Data might encompass sales, consumer preferences, and market trends related to essential consumer goods like food, beverages, and household products.

Alternative Factors: social responsible factors

Social responsibility factors are integral in evaluating the ethical and societal impacts of organizational operations.

These factors encompass:

- **Corporate Governance**: Assessment of company leadership, ethical conduct, and organizational transparency.
- **Employee Relations**: Examination of the relationships between management and employees, focusing on workplace conditions, employee satisfaction, and dispute resolutions.
- **Diversity and Inclusion**: Analysis of diversity within organizational structures, including board composition and equal opportunity practices.
- **Discrimination Practices**: Evaluation of company policies, actions, and preventative measures concerning discrimination and inequality.

Ticker	Community Contribution (CC)	Corporate Governance (CG)	Diversity (DIV)	Employee Relations (ER)	Environment (E)	Human Rights (HR)	Product (PRO)
AAPL	0.00	−25.00	50.00	0.00	14.29	0.00	−23.33
MSFT	0.00	25.00	100.00	5.56	35.71	50.00	10.00
AMZN	0.00	−25.00	−50.00	−16.67	0.00	0.00	−66.67
GOOGL	0.00	25.00	0.00	33.33	14.29	0.00	−23.33
TSLA	0.00	−50.00	0.00	−22.22	7.14	0.00	−16.67
FB	0.00	−25.00	50.00	0.00	0.00	50.00	−6.67
BRK.B	0.00	−25.00	0.00	0.00	7.14	0.00	0.00
JNJ	0.00	−25.00	0.00	44.44	14.29	0.00	−40.00
WMT	0.00	−25.00	−50.00	−16.67	14.29	0.00	−33.33
JPM	0.00	25.00	0.00	33.33	14.29	0.00	−16.67

Alternative Factors: social responsible factors

1. Community Contribution:
There is no fixed formula, but it's typically quantified through the monetary amount allocated to community development projects, sponsorships, and charitable donations, compared to the company's net income.

2. Corporate Governance:
It is evaluated through qualitative analysis of board structure, shareholder rights, executive compensation, and audit quality, among other criteria, occasionally leading to a governance score based on predefined criteria.

3. Diversity:
It can be quantified by calculating the diversity ratio or the percentage representation of various diversity categories within the organization, especially in leadership and decision-making roles.

4. Employee Relations:
Employee satisfaction surveys, turnover rates, and employee engagement levels are common metrics used to assess employee relations.

Alternative Factors: social responsible factors

5. Environment:
It is quantified using metrics like carbon emissions, water usage, waste production, and the implementation of environmentally friendly practices and policies.

6. Human Rights:
It's assessed through qualitative analysis of company policies, practices, and any reported incidents or violations related to human rights.

7. Product Responsibility:
It is typically assessed by evaluating product quality, safety standards, ethical sourcing, and marketing practices, and by monitoring product recalls, violations, and customer feedback.

Question/s of the week

Question: Compute the mean and variance of the random variable X = e^Z where Z is standard normal

Answer: sqrt(e) and e^2-e. For that solve for E(X) and Var(X)=E(X^2)-(EX)^2 using the integral for a continuance random variable.

Disclaimer

This course is for educational purposes only and does not offer investment advice or pre-packaged trading algorithms. The views expressed herein are not representative of any affiliated organizations or agencies. The main objective is to explore the specific challenges that arise when applying Data Science and Machine Learning techniques to financial data. Such challenges include, but are not limited to, issues like short historical data, non-stationarity, regime changes, and low signal-to-noise ratios, all of which contribute to the difficulty in achieving consistently robust results. The topics covered aim to provide a framework for making more informed investment decisions through a systematic and scientifically-grounded approach.