Masrik Dahir, Vivian Nguyen

Problem Statement

Data Acquisition and Technical Analysis for Blanket Assessments of the Stock Exchange (DATABASE): Artificial Intelligence to predict Equity Market Trend

DATABASE is an application that monitors the current state of the equity market, crypto market, and futures contract market and documents past transactions. DATABASE uses its vibrant database to find correlations and predict market trends.

Companies are entities that produce goods and services for their customers. Private companies and non-profit organizations can't issue shares of ownership like a stock. If a private company wants to go public, it has to offer shares for sale on an exchange (NYSE, NASDAQ, etc.). Every company has a founding date and a CEO. Public companies are operated by a board of directors and substantial shareholders vote on crucial decisions. Every public company has a stock symbol that is traded on a stock exchange. A publicly-traded company has an IPO (Initial Public Offering) date, which indicates the date when the company went public. Companies have to take short-term and long-term loans to buy equipment, services, and rent offices. Therefore, short-term and long-term loans are considered liabilities. The leverage ratio is the total liability and total shareholders' equity. The DATABASE database also categorizes each company by its industry. Each company has an EID (Employer ID) number. Internal Revenue Service (IRS) assigns a unique nine-digit number assigned to every U.S. business entity for identification.

User:

A user is an individual or identity. In the DATABASE database, the user entity contains basic attributes - mailing address, email address(es), phone number(s), and TIN. TIN is a unique identification number used by the IRS to identify identity or individual- the TIN of an entity is known as EIN and the TIN of an individual is known as Social Security Number (SSN).

Direct Investors:

Direct investors are entities or individuals who own stocks, options, mutual funds, and cryptocurrency from the stock exchanges or from the company itself. The direct investor can be

either an individual or an entity. Each direct Investor has its Taxpayer Identification Numbers (TIN). Each direct investor has a mailing name, address, email address, and phone number(s).

Broker:

Brokers buy and keep records of stocks for individuals and entities who are not direct investors. Brokers profit from the commission of each trade by its customers. Each broker has its unique EIN, name, address, phone number(s), and email address. Each broker must hold a stock-broker license to trade equities for its customers. Brokers can offer leverage trading. Leveraged trading is buying equities with borrowed money. The brokers charge interest on the borrowed funds. There is a margin of liquidation in leverage trading. If the leveraged value of the equity drops under a certain margin, the broker would liquidate the equity to protect the borrowed amount.

Exchanges:

In Exchanges, both brokers and traders buy and sell securities and cryptocurrencies. Every exchange has a market identifier code (MIC). The total market Cap of all securities in an exchange is the market cap of an exchange. Every exchange has a CEO, website, mailing address. An exchange may have multiple email addresses and phone numbers. Exchanges are operated by the currency of its country. Every exchange has Indices to track the performance of a group of assets in the exchange.

ETF:

An ETF is a security for a company. It is identified by the stock symbol. The ETF has some key attributes- market cap, volume, value, fully diluted shares. Market cap is the total value of a company's stock. Volume is the total value of stocks that are being actively traded. A fully diluted share is the total number of stocks of a company. Some ETFs offer quarterly dividends.

Option:

An option contract is a right (but not the obligation) to buy to sell a security/asset. An option has an expiration date until the contract holder can exercise his/her right to buy or sell the underlying

asset. The identifier of an option contract is the options symbol. Every options contract has a strike price.

Mutual Fund:

A mutual fund is a collection of securities held by an institution. The institution generally holds the right to buy or sell securities observing the market trend. Mutual funds are not listed on stock exchanges rather they are traded through financial professionals, brokerage firms, and directly from fund companies. A mutual fund's market opening and closing are also different from stock exchanges. Mutual funds are identified with stock symbols. Every mutual fund has net asset value, current market value, minimum investment. Mutual funds are categorized by asset class, the type of underlying assets. All mutual funds are required to distribute dividends.

Cryptocurrency:

Cryptocurrencies are identified by their currency symbols. Unlike stokes, cryptocurrencies have predictable inflation. The rate of inflation varies by currency. As a result, the current supply indicates current circulating coins. Total supply indicates the total number of coins in circulation including staked coins. Max supply indicates the maximum number of coins that can be in circulation in a lifetime. Every cryptocurrency has a market capitalization. Some cryptocurrency projects are open source (i.e., BTC, ETH, ADA, SOL, ALGO), while some projects are not open-source (i.e., HBAR). Unlike stocks, crypto-currencies can be exchanged in different exchanges including (centralized and decentralized exchange). The value for each coin is approximately the same on every exchange. The total value of stocks that are being actively traded in a particular exchange is the volume of the coin in the particular exchange.

Transaction:

The transaction is the transfer of a security or asset from one individual or entity to another individual or entity. Every transaction has a unit price. The transactions are concluded in an exchange. Each transaction must have the price and the quantity of the security/asset. The type of security/asset transferred is also recorded from each transaction.

Individual Investors:

Individual investors are also able to purchase and sell stocks through any financial service provider of their choice. One investor can use multiple services like Robinhood, Fidelity, Charles Schwab, etc. Individual investors can only have one account with each service but there can be multiple sub-accounts for each account. Individual investors are identified using their social security number and must have a name, date of birth, and address. An individual investor must provide one unique email address, although they are allowed to have anywhere from zero to multiple phone numbers.

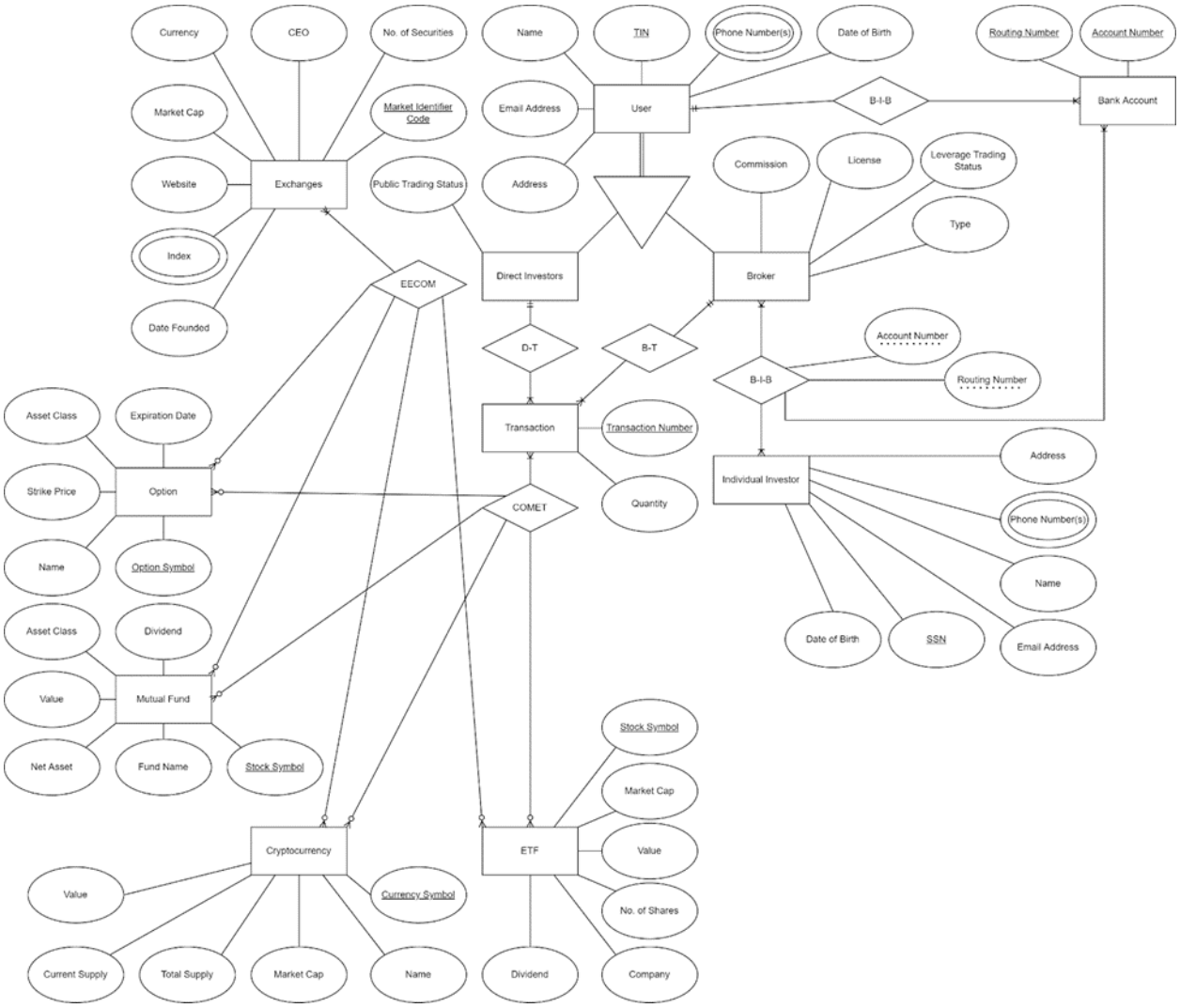
Brokerage Account:

Each user must have a brokerage account to trade securities and assets. Each brokerage account is linked to an individual's or entity's bank account. The investors are identified with their SSN numbers.

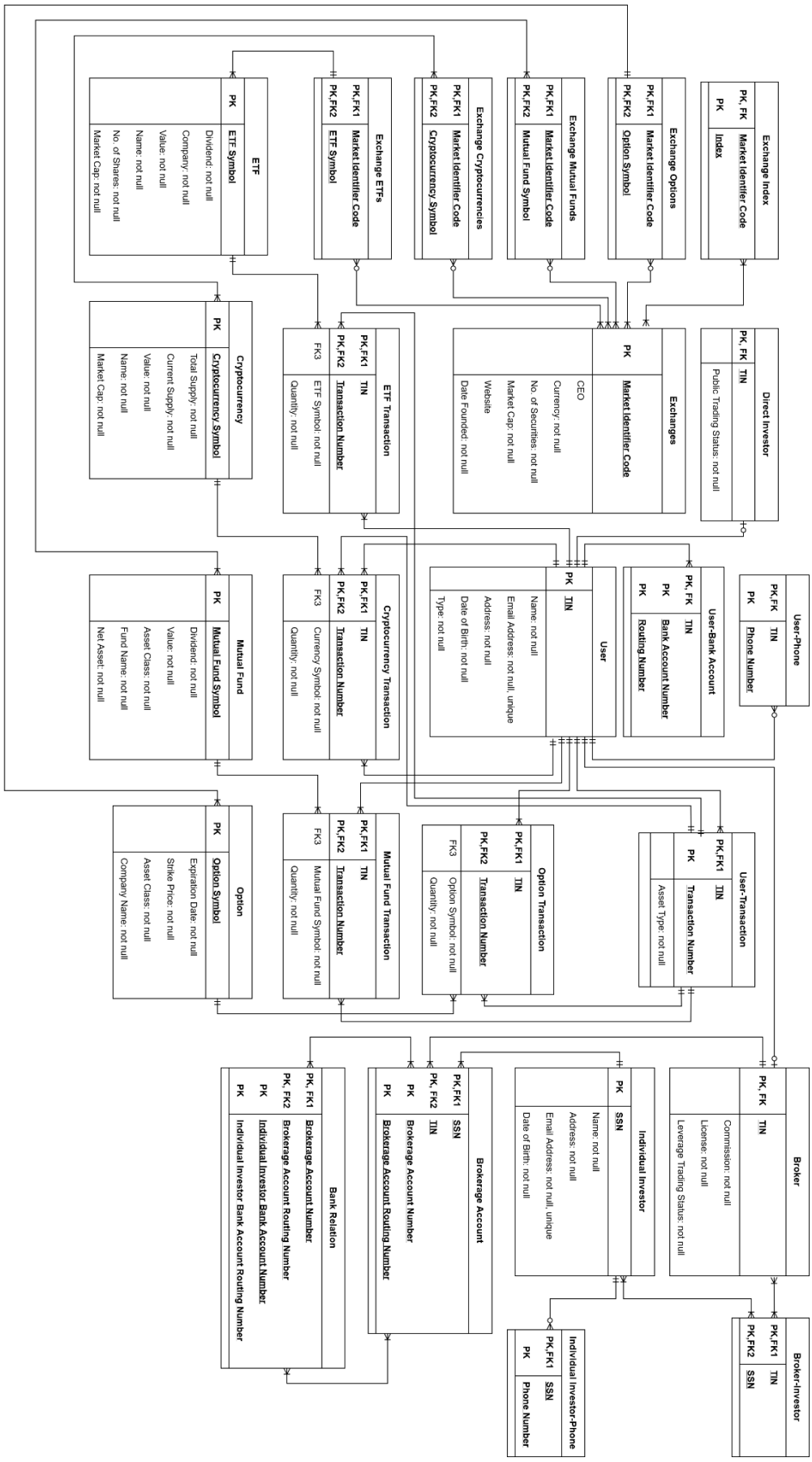
Bank Account:

Each user must have a bank account to transfer funds to a financial account. Bank accounts have an account number and a routing number.

Masrik Dahir, Vivian Nguyen



Masrik Dahir, Vivian Nguyen



Tables:

User (TIN: [PK], Name : [not null], Email Address : [not null, unique], Address : [not null], Date of Birth : [not null])

Broker (TIN: [PK], Commission : [not null], License : [not null], Leverage Trading Status : [not null])

FK {TIN} references {User.TIN}

Direct Investor (TIN: [PK], Public Trading Status : [not null])

FK {TIN} references {User.TIN}

Individual Investor (SSN: [PK], Address : [not null], Name : [not null], Email Address : [not null, unique], Date of Birth : [not null])

Option (Option Symbol: [PK], Expiration date : [not null], Strike Price : [not null], Asset Class : [not null], Name : [not null])

Mutual Fund (Mutual Fund Symbol: [PK], Dividend : [not null], Value : [not null], Asset Class : [not null], Fund Name : [not null], Net Asset : [not null])

Cryptocurrency (Cryptocurrency Symbol: [PK], Total Supply : [not null], Current Supply : [not null], Value : [not null], Name : [not null], Market Cap : [not null])

ETF (ETF Symbol: [PK], Dividend : [not null], Company : [not null], Value : [not null], Name : [not null], No. of Shares : [not null], Market Cap : [not null])

Exchange (Market Identifier Code: [PK], CEO, Currency : [not null], No. of Securities : [not null], Market Cap : [not null], Website, Date Founded : [not null])

Option Transaction (TIN: [PK], Transaction Number: [PK], Option Symbol : [not null], Quantity : [not null])

FK {TIN} references {User.TIN}

FK {Transaction Number} references {User Transaction.Transaction Number}

FK {Option Symbol} references {Option.Option Symbol}

Mutual Fund Transaction (TIN: [PK], Transaction Number: [PK], Mutual Fund Stock Symbol : [not null], Quantity : [not null])

FK {TIN} references {User.TIN}

FK {Transaction Number} references {User Transaction.Transaction Number}

FK {Mutual Fund Symbol} references {Mutual Fund.Mutual Fund Symbol}

Cryptocurrency Transaction (TIN: [PK], Transaction Number: [PK], Cryptocurrency Symbol : [not null], Quantity : [not null])

FK {TIN} references {User.TIN}

FK {Transaction Number} references {User Transaction.Transaction Number}

FK {Cryptocurrency Symbol} references {Cryptocurrency.Cryptocurrency Symbol}

ETF Transaction (TIN: [PK], Transaction Number: [PK], ETF Stock Symbol : [not null], Quantity : [not null])

FK {TIN} references {User.TIN}

FK {Transaction Number} references {User Transaction.Transaction Number}

FK {ETF Symbol} references {ETF.ETF Symbol}

User Bank Account (TIN: [PK], Routing Number: [PK], Account Number: [PK])

FK {TIN} references {User.TIN}

User Phone (TIN: [PK], Phone Number: [PK])

FK {TIN} references {User.TIN}

Broker Investor (TIN: [PK], SSN: [PK])

FK {TIN} references {User.TIN}

FK {SSN} references {Individual Investor.SSN}

Individual Investor Phone (SSN: [PK], Phone Number: [PK])

FK {SSN} references {Individual Investor.SSN}

Exchange Index (Market Identifier Code: [PK], Index: [PK])

FK {Market Identifier Code} references {Exchanges.Market Identifier Code}

Exchange Options (Market Identifier Code: [PK], Option Symbol: [PK])

FK {Market Identifier Code} references {Exchanges.Market Identifier Code}

FK {Option Symbol} references {Option.Option Symbol}

Exchange Mutual Funds (Market Identifier Code: [PK], Mutual Fund Stock Symbol: [PK])

FK {Market Identifier Code} references {Exchanges.Market Identifier Code}

FK {Mutual Fund Symbol} references {Mutual Fund.Mutual Fund Symbol}

Exchange ETFs (Market Identifier Code: [PK], ETF Stock Symbol: [PK])

FK {Market Identifier Code} references {Exchanges.Market Identifier Code}

FK {ETF Symbol} references {ETF.ETF Symbol}

Exchange Cryptocurrencies (Market Identifier Code: [PK], Cryptocurrency Symbol: [PK])

FK {Market Identifier Code} references {Exchanges.Market Identifier Code}

FK {Cryptocurrency Symbol} references {Cryptocurrency.Cryptocurrency Symbol}

User Transaction (TIN: [PK], Transaction Number: [PK])

FK {TIN} references {User.TIN}

Bank Relation (Brokerage Account Number: [PK], Brokerage Account Routing Number: [PK], Bank Account Number: [PK], Bank Routing Number: [PK])

FK {Brokerage Account Number} references {Brokerage Account.Brokerage Account Number}

FK {Brokerage Account Routing Number} references {Brokerage Account.Brokerage Account Routing Number}

Brokerage Account (TIN: [PK], SSN: [PK], Brokerage Account Number: [PK], Brokerage Routing Number: [PK])

FK {TIN} references {Broker.TIN}

FK {SSN} references {Individual Investor.SSN}

Enumeration of tables with their functional dependencies and keys:

User (TIN, Name, Email Address, Address, Date of Birth)

F = { TIN → Name, Email Address, Address, Date of Birth }

PK = { TIN }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Broker (Broker TIN, Commission, License, Leverage Trading Status)

F = { Broker TIN → Commission, License, Leverage Trading Status }

PK = { TIN }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Direct Investor (Direct Investor TIN, Public Trading Status)

F = { TIN → Public Trading Status }

PK = { TIN }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Individual Investor (SSN, Address, Name, Email Address, Date of Birth)

F = { SSN → Address, Name, Email Address, Date of Birth }

PK = { SSN }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Option (Option Symbol, Expiration date, Strike Price, Asset Class, Name)

F = { Option Symbol→ Expiration date, Strike Price, Asset Class, Name }

PK = { Option Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Mutual Fund (Mutual Fund Symbol, Dividend, Value, Asset Class, Fund Name, Net Asset)

F = { Mutual Fund Symbol→ Dividend, Value, Asset Class, Fund Name, Net Asset }

PK = { Mutual Fund Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Cryptocurrency (Cryptocurrency Symbol, Total Supply, Current Supply, Value, Name, Market Cap)

F = { Cryptocurrency Symbol→ Total Supply, Current Supply, Value, Name, Market Cap }

PK = { Cryptocurrency Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

ETF (ETF Symbol, Dividend, Company, Value, Name, No. of Shares, Market Cap)

F = { ETF Symbol→ Dividend, Company, Value, Name, No. of Shares, Market Cap }

PK = { ETF Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Exchange (Market Identifier Code, CEO, Currency, No. of Securities, Market Cap, Website, Date Founded)

F = { Market Identifier Code → CEO, Currency, No. of Securities, Market Cap, Website, Date Founded, Currency Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Option Transaction Brokerage Account (TIN, Individual SSN, Brokerage Account Number, Brokerage Routing Number)

F = { TIN, Transaction Number → Option Symbol, Quantity }

PK = { TIN, Transaction Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Mutual Fund Transaction (TIN, Transaction Number, Mutual Fund Stock Symbol, Quantity)

F = { TIN, Transaction Number → Mutual Fund Stock Symbol, Quantity }

PK = { TIN, Transaction Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

Cryptocurrency Transaction (TIN, Transaction Number, Cryptocurrency Symbol, Quantity)

F = { TIN, Transaction Number → Currency Symbol, Quantity }

PK = { TIN, Transaction Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

ETF Transaction (TIN, Transaction Number, ETF Stock Symbol, Quantity)

F = { TIN, Transaction Number → ETF Stock Symbol, Quantity }

PK = { TIN, Transaction Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and there are no multivalued dependencies.

User Bank Account (TIN, Routing Number, Account Number)

F = { TIN →→ Routing Number, Account Number }

PK = { TIN, Routing Number, Account Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

User Phone (TIN, Phone Number)

F = { TIN →→ Phone Number }

PK = { TIN, Phone Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Broker Investor (Broker TIN, SSN)

F = { TIN →→ SSN }

PK = { TIN, SSN }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Individual Investor Phone (SSN, Phone Number)

F = { SSN →→ Phone Number }

PK = { SSN →→ Phone Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Exchange Index (Market Identifier Code, Index)

F = { Market Identifier Code →→ Index }

PK = { Market Identifier Code, Index }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Exchange Options (Market Identifier Code, Option Symbol)

F = { Market Identifier Code →→ Option Symbol }

PK = { Market Identifier Code, Option Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Exchange Mutual Funds (Market Identifier Code, Mutual Fund Stock Symbol)

F = { Market Identifier Code →→ Mutual Fund Stock Symbol }

PK = { Market Identifier Code, Mutual Fund Stock Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Exchange ETFs (Market Identifier Code, ETF Stock Symbol)

F = { Market Identifier Code →→ ETF Stock Symbol }

PK = { Market Identifier Code, ETF Stock Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Exchange Cryptocurrencies (Market Identifier Code, Cryptocurrency Symbol)

F = { Market Identifier Code →→ Cryptocurrency Symbol }

PK = { Market Identifier Code, Cryptocurrency Symbol }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

User Transaction (TIN, Transaction Number)

F = { TIN →→ Transaction Number }

PK = { TIN, Transaction Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Bank Relation (Brokerage Account Number, Brokerage Account Routing Number, Bank Account Number, Bank Routing Number)

F = { Brokerage Account Number, Brokerage Account Routing Number →→ Bank Account Number, Bank Routing Number }

PK = { Brokerage Account Number, Brokerage Account Routing Number, Bank Account Number, Bank Routing Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.

Brokerage Account (Brokerage Account Number, Brokerage Routing Number)

F = { TIN, Individual SSN →→ Account Number, Routing Number }

PK = { TIN, Individual SSN, Account Number, Routing Number }

It is in 4NF because the antecedent of the FD is the candidate key of the relationship and the antecedent and consequents together form the whole set of attributes of the relationship.