Project 1

Group 31: Saumya Gupta (2016ME10689), Rohit Kumar Singh (2016ME10080)

Due date: March 2, 2020, 11:55pm IST

1 Project Description

- **Motivation**: This project is a database of the full historical daily price and volume data for all US-based stocks trading on the NYSE, NASDAQ, and AMEX US market exchanges till 11/10/2017. This data has been used to introduce a platform to analyse the database and use it for further transactions. For the purpose of demonstration, each profile is introduced with a certain amount as balance to start from which can be used to perform transactions on the platform. Further, the platform also contains data regarding the company such as Industry, Sector and a brief description about the company. The data is sorted in various ways such as Sector based distribution, Industry based distribution so that the user can also see the general trend of the Industry or the Sector.

- **Stack (Frontend & Backend) Description**:
  For the full stack development we have used PostgresSQL12+flask (python3 based microframework) + Angular (Typescript based frontend framework).

  We used flask for its simplicity in communicating with database server and frontend server along with python's superior data handling.Python was overall easy to manipulate queries which otherwise would have been difficult to write.
  For the frontend we used Angular as it provided modular approach and flexible and reliable design support.Angular also provide us with services which makes it easy to develop things such as authguard and http communications with the backend server.

- **Entities & Relations**:
  Find the ER diagram on the next page. Following is the list of entities and relations in our model.

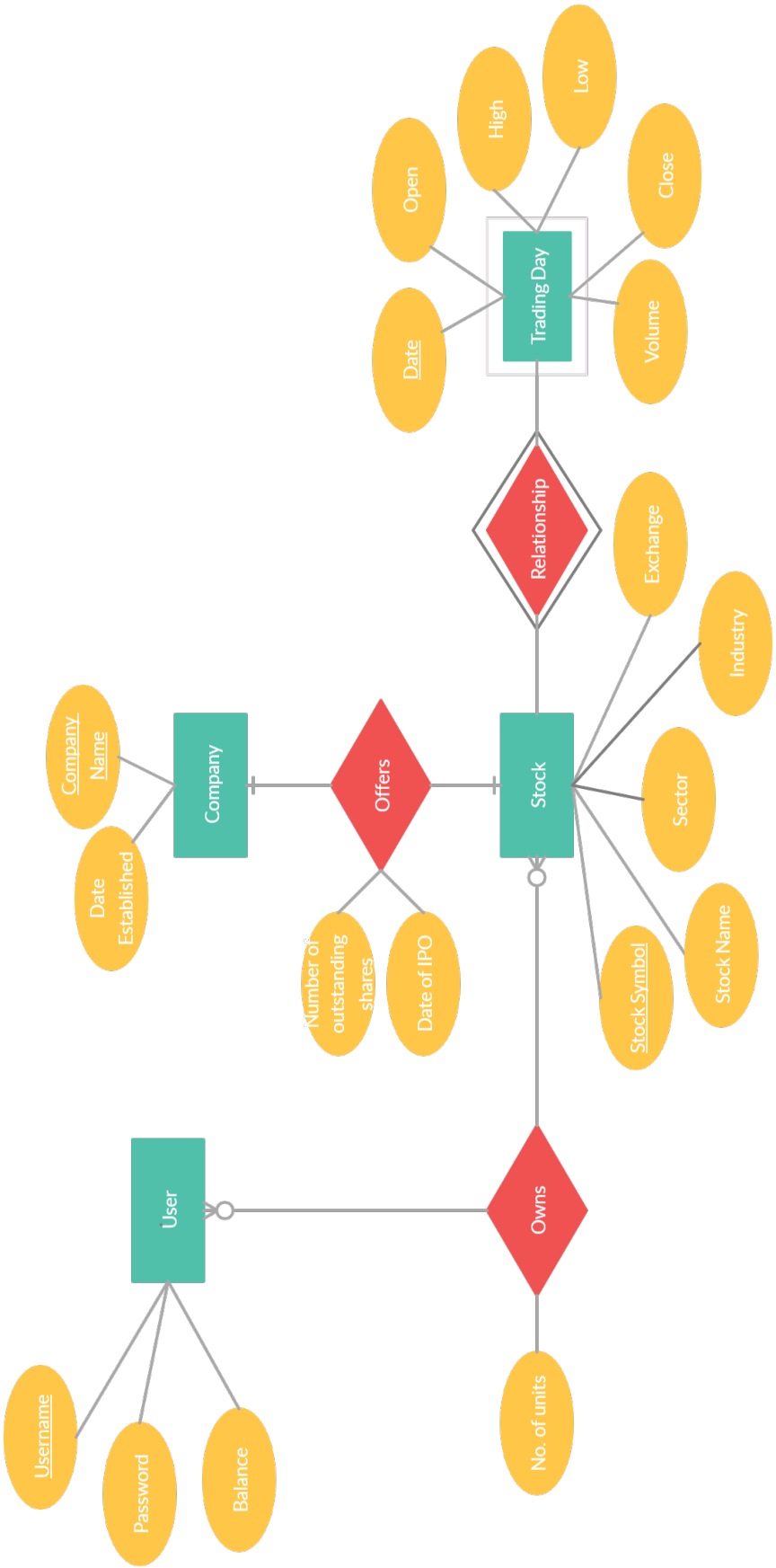
Table 1: Entities and Attributes table

Entities	**Primary Key**	**Other Attributes**
User	UserName,	Password, Balance
Stock	Stock Symbol	Name, Sector, Industry, Exchange
Company	Company Name	Date Established
Trading Day (Weak Entity - Depends on Stock)	Date	Open, High, Low, Close, Volume

Table 2: Relationships and Attributes table

Relationships	**Related Entity Sets**	**Attributes**
Owns	User Owns Stock	Stock Units
Offers	Company Offers Stock	Name, Sector, Industry, Exchange
Traded on	Stock is Traded on Trading Day	*None*

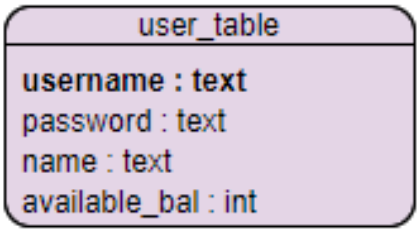
Figure 1: Entity Relationship Diagram

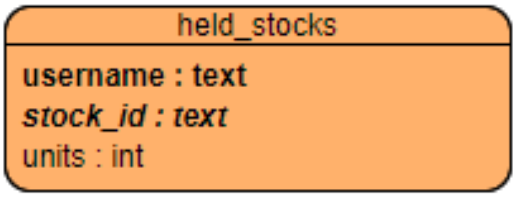


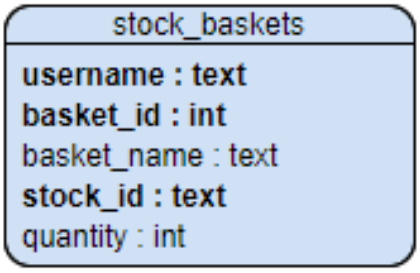
2 Data Description

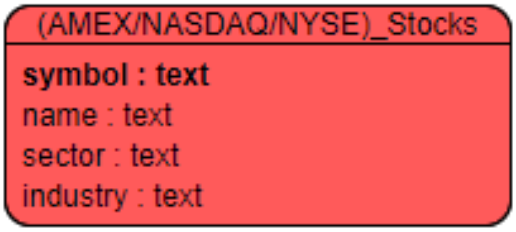
1. The source of this data set is the Kaggle website, link for the same is (https://www.kaggle.com/borismarjanovic/price-volume-data-for-all-us-stocks-etfs).
2. The data was in "Ready-made" format on the aforementioned website.
3. There are a total of 7,198 tables that have been taken from the above data set and have been added to the database of this project. Tables have an average of 3000 tuples each. The total size of the data is 670 MB.

Description of tables in the database









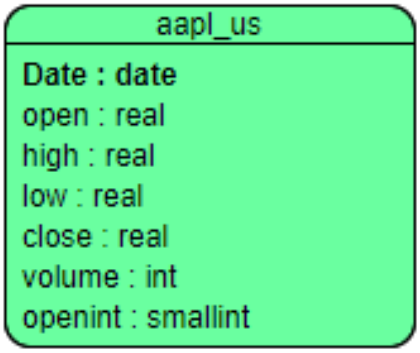


Table 3: Tables and loading times

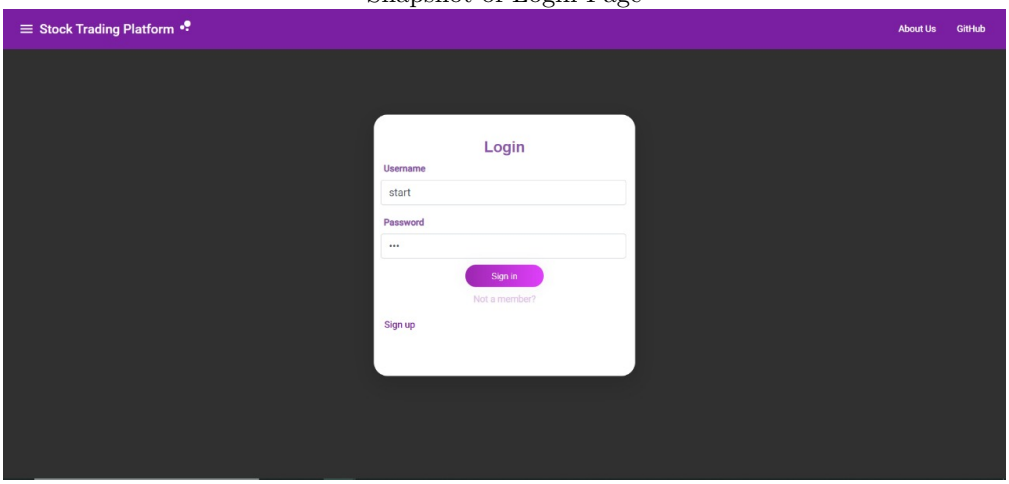
Tables	Number of tables	Number of Tuples per table	Time to Load
Stock Tables	7198 tables for 7198 stocks	∼3000 records	∼40 mins for all tables ∼500 ms for each table
Exchange Tables	3 tables for 3 exchanges	∼3500 records	∼1500 ms for all tables
User Table	1 table	Initially 0 records, added with time	None

3 Functionality & UI

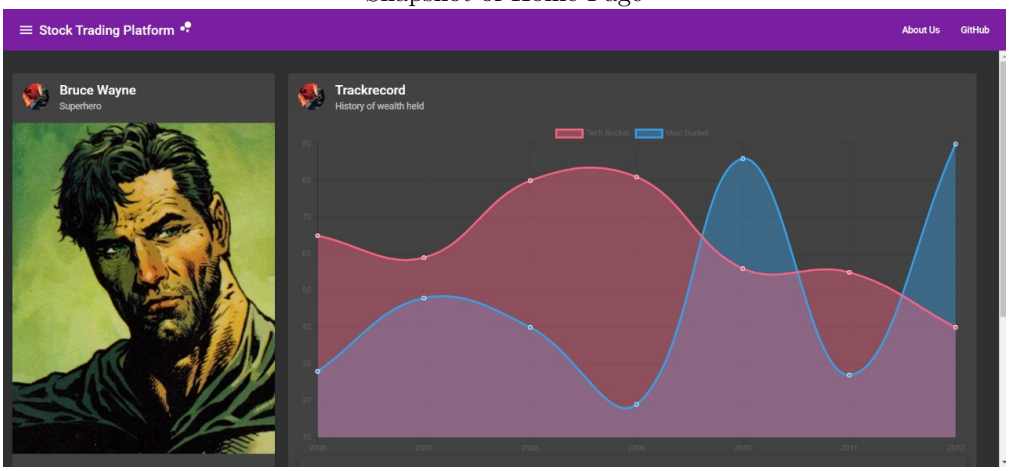
1. User's View of the System

   (a) Home: This will contain the userdetails and other relevant thing

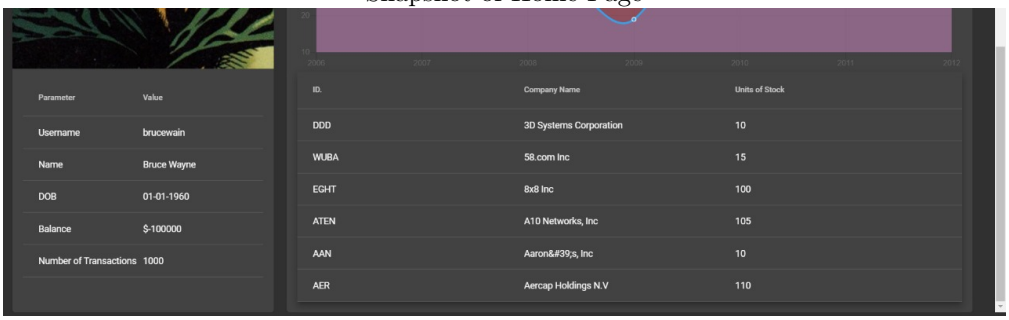
Snapshot of Login Page



Snapshot of Home Page

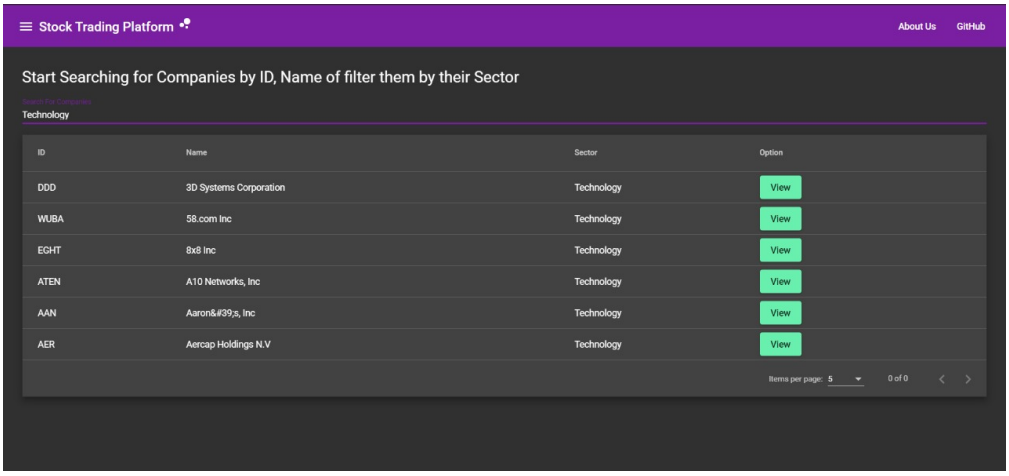


Snapshot of Home Page



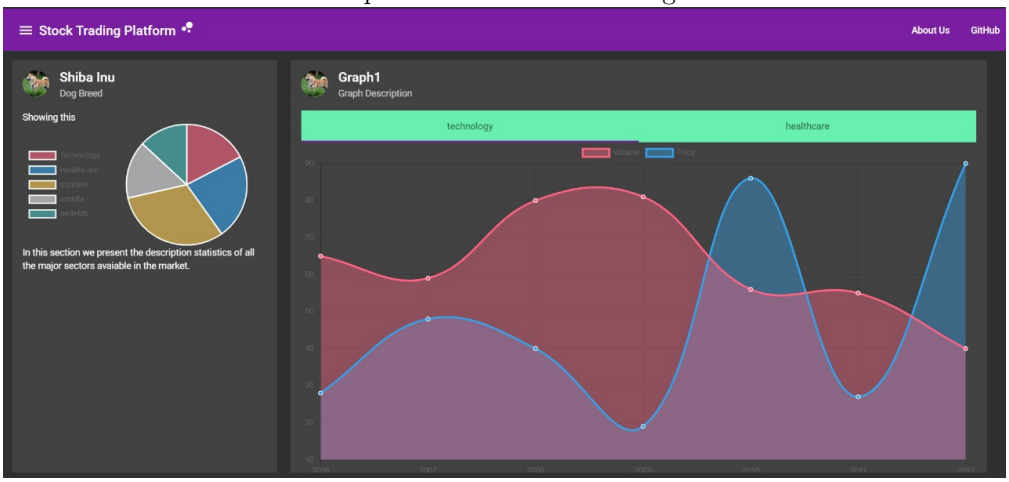
   (b) Stock Profile: This will contain the list of all stocks and we can view details of each stock by selecting it on it

Snapshot of Stock Profile Page

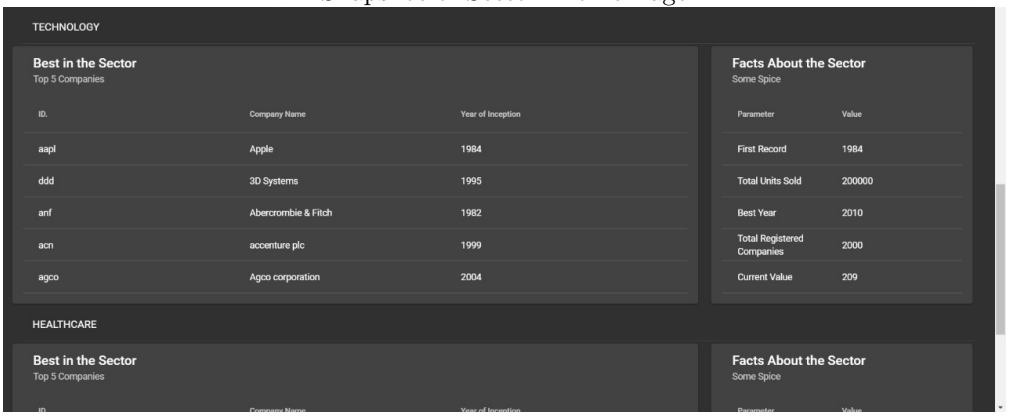


(c) Sector Profile: Here we present some statistics about each sector available and overall stats about the sector.

Snapshot of Sector Profile Page

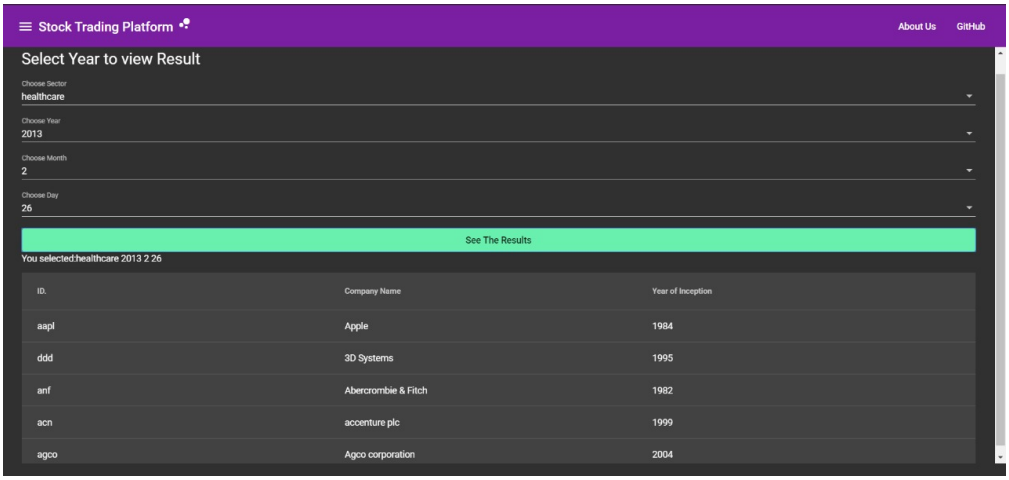


Snapshot of Sector Profile Page



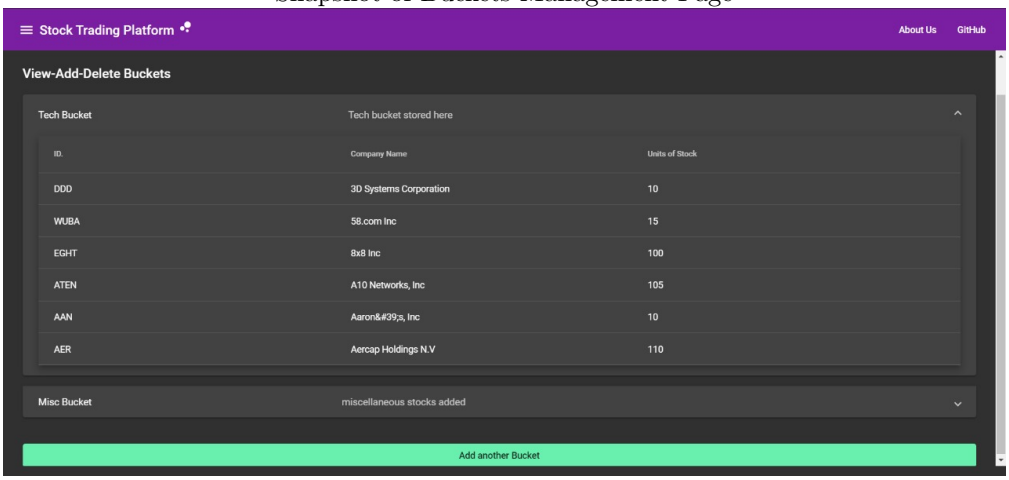
(d) Analytics: Here we can perform many query to get the results such as best stocks on certain date and year.

Snapshot of Analytics Page



(e) My Bucket : Here user can make his bucket see its performance in the market.

Snapshot of Buckets Management Page



2. Special Functionality

   (a) Indexes: The list of companies has over 7,000 rows, so we indexed the table on Symbol for fast lookups of a companies. In addition, we also indexed the Dates for each company table, to display the prices of the stock of a particular company at the browser end quickly.

   (b) Constraints: The primary key for the Stock list(list of companies) is the Symbol of the company. The primary key for each of the companies stock prices is the Date of that particular entry. For the user table it refers to is the username. For the held stocks table it refers to the tuple (username, stock id) and finally for the basket table it refers to the tuple (username, basket id, stock id).

3. List of queries

   (a) Login Page:

      i. SELECT password FROM users WHERE username = 1;

   (b) Stock Profile

      i. The first query will be used to retrieve the symbol of the Stock from the list of all Companies. E.g. SELECT symbol from NASDAQ_Stocks WHERE name = 'Apple Inc.'; We are using regex to perform patter matching.

      ii. Next we will use python along with psycopg2 library to retrieve the table of Stock prices for that particular company on various dates.

      iii. Next query will be used to select the start date and the end date to finally make the graph of the variation of Stock prices versus the date. E.g.
      cur.execute("SELECT Symbol FROM NASDAQ_Stocks where name = 'Apple Inc.'")

one = cur.fetchone()
cur.execute("SELECT * FROM "+one[0]+"_us WHERE Date ≥ "+start_date+" and Date ≤"+end_date+";")
all = cur.fetchall()

iv. The user can further select a particular date from the graph to get prices on individual dates.

(c) Sector Profile

i. The first query will be used to retrieve the symbol of all the Stock belonging to a particular sector from the list of all Companies.
E.g. SELECT symbol from NASDAQ_Stocks where Sector = 'Technology';

ii. Similar to the Stock profile dates we can select a start and end date to view the graph between those dates.

iii. Overall growth of the sector can be observed by summing up the stock prices of all the companies(Symbols).

iv. A pie chart showing the distribution of the companies in various charts will also be available.

(d) Analytics:

i. To view the percentage return of a particular company. E.g.
SELECT 100.0*(curr.Open - prev.Open) / prev.Open As PercentDiff
FROM aapl_us As curr
JOIN aapl_us As prev
ON curr.Date = '2017-10-24' AND prev.Date = '2016-10-24' limit 1;

ii. A list of the top 5/10/20 companies based on returns can be seen by calculating the percentage returns for all the companies between certain dates.

(e) My Basket

i. Here the user can make a basket by selecting Stocks and specifying the units of the stocks.

ii. Then again the tables of each of the Stock in the basket will be retrieved.

iii. After the tables are retrieved the Stock prices will be summed up.

iv. The summed up Stock prices will be used to represent the variation of the stock prices over the days.