Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

I. Preparing DATA

Step1. Importing DATA

1.To complete this step, we need the following:

A. Go to google workspace marketplace
B. Open a new spreadsheet and under the Add-ons menu choose “Get Add-ons.”
C. This will bring you to the Google Workspace Marketplace. Search for example “Alpha Vantage Market Data” and click through.
D. When you should see “Alpha Vantage Market Data” add-on, click the “Install” button to install the Add-on. Google will request that you approve permissions necessary for the Addon to function.
E. click “Allow” to enable the Add-on. If successful you should see a confirmation screen. Refresh your browser window to make sure the Add-on menu is initialized. Then enter your API Key which will open a dialog where you can enter your API Key. You’re ready to go!

2. having “Alpha vantage..” add-on installed, in a spreadsheet use this following script to import data for six stocks, for the six last days. choose 60min as interval of daytime :

- - - - - - - - - - - -

=IMPORTDATA("https://www.alphavantage.co/query?function=TIME_SERIES_INTRADAY&symbol=TSLA&interval=60min&apikey=AIzaSyDyor2KbXMksDSdt1GDHRLMeAgoFNBssg0")

- - - - - - - - - - - - - - - - -

Step2 : Extract the relevant DATA from the spreadsheet and reorganize your data to store it in SQL DB

(three day time for each of the six days , for example : 10:00; 14:00; 18:00)

for example for apple stock, this is the final look of DATA in google spreadsheet :

	A	B	C	D	E	F	G
1	AAPL	day1	day2	day3	day4	day5	day6
2		20-10-2023	23-10-2023	24-10-2023	25-10-2023	26-10-2023	27-10-2023
3	18:00	172,71	172,87	173,18	169,77	166,47	168,43
4	14:00	173,51	173,5	173,12	171,09	167,68	167,62
5	10:00	173,36	172,21	172,53	172,14	166,77	167,98

Step3 : Create table in SQL using DB Browser for SQLite

The code :

```
CREATE TABLE stocks (
  id INTEGER NOT NULL PRIMARY KEY,
  name TEXT,
  symbol TEXT,
  price REAL,
datetime TEXT
);
```

Here is the DATA structure we get when run the code :

Tables (1)

Name	Type	Schema
stocks		CREATE TABLE "stocks"(symbol TEXT, name TEXT, datetime TEXT, price REAL)
symbol	TEXT	"symbol" TEXT
name	TEXT	"name" TEXT
datetime	TEXT	"datetime" TEXT
price	REAL	"price" REAL

Step4 : insert new records in table

Code :

```
BEGIN;
INSERT INTO stocks (id, name, symbol, price, datetime)
VALUES
   (1, 'Apple', 'AAPL', 173.36, '2023-11-20 10:00:00'),
   (2, 'Apple', 'AAPL',  280.30, '2023-11-20 14:00:00'),
   (3, 'Apple', 'AAPL',  2750.00, '2023-11-20 18:00:00')
    - - more rows
COMMIT;
```

To insert new records fast and more efficiently, try to chunk data and insert it using separate statement for each stock as mentionned in the code above.

The following is how the table looks like when you run this code (to display data for 2 stocks : apple and amazon :

```
SELECT (symbol, name, datetime, price)
FROM stocks
LIMIT 36;
```

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

symbol	name	datetime	price
AAPL	Apple	2023-10-23 14:00:00	173.5
AAPL	Apple	2023-10-23 18:00:00	172.86
AAPL	Apple	2023-10-24 10:00:00	172.21
AAPL	Apple	2023-10-24 10:00:00	172.53
AAPL	Apple	2023-10-24 14:00:00	173.12
AAPL	Apple	2023-10-24 18:00:00	173.18
AAPL	Apple	2023-10-25 10:00:00	172.16
AAPL	Apple	2023-10-25 14:00:00	171.1
AAPL	Apple	2023-10-25 18:00:00	169.77
AAPL	Apple	2023-10-26 10:00:00	168.32
AAPL	Apple	2023-10-26 14:00:00	167.69
AAPL	Apple	2023-10-26 18:00:00	166.47
AAPL	Apple	2023-10-27 10:00:00	168.32
AAPL	Apple	2023-10-27 14:00:00	167.62
AAPL	Apple	2023-10-27 18:00:00	167.99
AAPL	Apple	2023-10-30 10:00:00	169.44
AAPL	Apple	2023-10-30 14:00:00	170.5
AAPL	Apple	2023-10-30 18:00:00	164.87
AMZN	Amazon	2023-10-23 10:00:00	125.76
AMZN	Amazon	2023-10-23 14:00:00	127.62
AMZN	Amazon	2023-10-23 18:00:00	126.93
AMZN	Amazon	2023-10-24 10:00:00	128.35
AMZN	Amazon	2023-10-24 14:00:00	127.77
AMZN	Amazon	2023-10-24 18:00:00	126.8
AMZN	Amazon	2023-10-25 10:00:00	123.22
AMZN	Amazon	2023-10-25 14:00:00	121.58
AMZN	Amazon	2023-10-25 18:00:00	119.85
AMZN	Amazon	2023-10-26 10:00:00	160.76
AMZN	Amazon	2023-10-26 14:00:00	120.46
AMZN	Amazon	2023-10-26 18:00:00	125.38
AMZN	Amazon	2023-10-27 10:00:00	129.46
AMZN	Amazon	2023-10-27 14:00:00	127.99
AMZN	Amazon	2023-10-27 18:00:00	127.76
AMZN	Amazon	2023-10-30 10:00:00	130.63
AMZN	Amazon	2023-10-30 14:00:00	132.64
AMZN	Amazon	2023-10-30 18:00:00	132.43

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

II. Performing basic analysis on the data and identifying trends.

1. What are the distinct stocks in the table?

   code :

   ```
   SELECT DISTINCT name, symbol
   FROM stocks;
   ```

   The following table shows the distinct stocks and their symbol :

name	symbol
Apple	**AAPL**
Amazon	**AMZN**
Google	**GOOGL**
Metaverse	**META**
Microsoft	**MSFT**
Tesla	**TSLA**

2. Querying all data for a single stock. Is there any overall trends?

   The Code to query all data for Apple (AAPL) as example:

   ```
   SELECT *
   FROM stocks
   WHERE symbol = 'AAPL' ;
   ```

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

The result set :

symbol	name	datetime	price
AAPL	Apple	2023-10-23 14:00:00	173.5
AAPL	Apple	2023-10-23 18:00:00	172.86
AAPL	Apple	2023-10-23 10:00:00	172.21
AAPL	Apple	2023-10-24 10:00:00	172.53
AAPL	Apple	2023-10-24 14:00:00	173.12
AAPL	Apple	2023-10-24 18:00:00	173.18
AAPL	Apple	2023-10-25 10:00:00	172.16
AAPL	Apple	2023-10-25 14:00:00	171.1
AAPL	Apple	2023-10-25 18:00:00	169.77
AAPL	Apple	2023-10-26 10:00:00	168.32
AAPL	Apple	2023-10-26 14:00:00	167.69
AAPL	Apple	2023-10-26 18:00:00	166.47
AAPL	Apple	2023-10-27 10:00:00	168.32
AAPL	Apple	2023-10-27 14:00:00	167.62
AAPL	Apple	2023-10-27 18:00:00	167.99
AAPL	Apple	2023-10-30 10:00:00	169.44
AAPL	Apple	2023-10-30 14:00:00	170.5
AAPL	Apple	2023-10-30 18:00:00	164.87

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

Trend analysis :

Based on the provided data for the AAPL (Apple Inc.) stock prices at different times over several days, we can observe the following trends:

- On October 23rd, the price started at $172.21, peaked at $173.5, and then dipped slightly to $172.86 by the end of the day. This indicates a relatively stable day with a slight increase in price from the opening to the closing.
- On October 24th, the price showed a small upward trend throughout the day, starting at $172.53 and closing slightly higher at $173.18.
- However, starting on October 25th, there appears to be a downward trend. The price opened at $172.16 and fell to $169.77 by the end of the day.
- This downward trend continued on October 26th, with the price starting at $168.32 and decreasing further to $166.47 by the end of the day.
- On October 27th, the price fluctuated around the same values as it opened ($168.32) and closed ($167.99) with a slight dip in between.
- On October 30th, there was a significant increase from the opening price of $169.44 to a peak of $170.5, but then a sharp decline to $164.87 by the end of the day.

Overall, if we look at the opening price on October 23rd ($172.21) and compare it to the closing price on October 30th ($164.87), there is a noticeable overall downward trend in the stock price over this week-long period.

3. Which rows have a price above a certain value (for example : the average)

The code :

```
WITH average_price AS (
  SELECT name, AVG(price) AS avgp
  FROM stocks
  GROUP BY name
)
SELECT s.name, s.symbol, s.datetime, s.price
FROM stocks s
JOIN average_price ap ON s.name = ap.name
```

```
WHERE s.price > ap.avgp AND s.symbol = 'AAPL';
```

The result set for AAPL as example :

name	symbol	datetime	price
Apple	**AAPL**	**2023-10-23 14:00:00**	**173.5**
Apple	**AAPL**	**2023-10-23 18:00:00**	**172.86**
Apple	**AAPL**	**2023-10-24 10:00:00**	**172.21**
Apple	**AAPL**	**2023-10-24 10:00:00**	**172.53**
Apple	**AAPL**	**2023-10-24 14:00:00**	**173.12**
Apple	**AAPL**	**2023-10-24 18:00:00**	**173.18**
Apple	**AAPL**	**2023-10-25 10:00:00**	**172.16**
Apple	**AAPL**	**2023-10-25 14:00:00**	**171.1**
Apple	**AAPL**	**2023-10-30 14:00:00**	**170.5**

4. Sorting the table by price. What are the minimum and maximum prices?

The code :

```
SELECT *
FROM stocks
WHERE symbol = 'AAPL'
ORDER by price DESC;
```

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

The result set :

name	symbol	datetime	price
Apple	AAPL	2023-10-23 14:00:00	173.5
Apple	AAPL	2023-10-24 18:00:00	173.18
Apple	AAPL	2023-10-24 14:00:00	173.12
Apple	AAPL	2023-10-23 18:00:00	172.86
Apple	AAPL	2023-10-24 10:00:00	172.53
Apple	AAPL	2023-10-24 10:00:00	172.21
Apple	AAPL	2023-10-25 10:00:00	172.16
Apple	AAPL	2023-10-25 14:00:00	171.1
Apple	AAPL	2023-10-30 14:00:00	170.5
Apple	AAPL	2023-10-25 18:00:00	169.77
Apple	AAPL	2023-10-30 10:00:00	169.44
Apple	AAPL	2023-10-26 10:00:00	168.32
Apple	AAPL	2023-10-27 10:00:00	168.32
Apple	AAPL	2023-10-27 18:00:00	167.99
Apple	AAPL	2023-10-26 14:00:00	167.69
Apple	AAPL	2023-10-27 14:00:00	167.62
Apple	AAPL	2023-10-26 18:00:00	166.47
Apple	AAPL	2023-10-30 18:00:00	164.87

For Aplle inc. MIN(price) = 164.87 and MAX(price) = 173.5

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

III. Analyzing DATA using aggregates - - - -

1. Exploring using aggregate functions to look at key statistics about the data (e.g., min, max, average) :

The code :

```
WITH aggregates_temp AS (
  SELECT
        name,
        ROUND(AVG(price), 2) AS avgp,  -- Assuming 2 decimal places for
rounding
        MAX(price) AS maxp,
         MIN(price) AS minp,
        (MAX(price) - MIN(price)) AS spread
  FROM stocks
  GROUP BY name
)
SELECT
  at.name,
  at.avgp,
  at.maxp,
  at.minp,
  at.spread
FROM aggregates_temp at;
```

The result set :

name	avgp	maxp	minp	spread
Amazon	128.63	160.76	119.85	40.91
Apple	170.09	173.5	164.87	8.63
Google	128.16	139.03	121.22	17.81
Metaverse	302.63	317.36	289.65	27.71
Microsoft	333.97	343.23	328.28	14.95
Tesla	209.67	217.86	196.29	21.57

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

Overall Analysis:

- ➔ **Volatility:** Amazon's stock shows the most volatility with the largest spread, while Apple's stock appears to be the most stable with the smallest spread.
- ➔ **Price Levels:** Microsoft has the highest average stock price, which could be indicative of its market valuation in comparison to the others.
- ➔ **Market Fluctuations:** The differences in spread could be due to various factors such as market sentiment, company performance, industry changes, or broader economic events.

To perform more detailed insights, we would typically need to consider additional context such as industry performance, company news, and broader market conditions during this period.

2. Grouping the data by day or hour of day. Does day of week or time of day impact prices?

Code :

```
SELECT *
FROM stocks
GROUP BY name, datetime -- grouping by day
ORDER BY symbol ASC;
```

Result set :

symbol	name	datetime	price
AAPL	Apple	2023-10-23 14:00:00	173.5
AAPL	Apple	2023-10-23 18:00:00	172.86
AAPL	Apple	2023-10-23 10:00:00	172.21
AAPL	Apple	2023-10-24 14:00:00	173.12
AAPL	Apple	2023-10-24 18:00:00	173.18
AAPL	Apple	2023-10-25 10:00:00	172.16

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

AAPL	Apple	2023-10-25 14:00:00	171.1
AAPL	Apple	2023-10-25 18:00:00	169.77
AAPL	Apple	2023-10-26 10:00:00	168.32
AAPL	Apple	2023-10-26 14:00:00	167.69
AAPL	Apple	2023-10-26 18:00:00	166.47
AAPL	Apple	2023-10-27 10:00:00	168.32
AAPL	Apple	2023-10-27 14:00:00	167.62
AAPL	Apple	2023-10-27 18:00:00	167.99
AAPL	Apple	2023-10-30 10:00:00	169.44
AAPL	Apple	2023-10-30 14:00:00	170.5
AAPL	Apple	2023-10-30 18:00:00	164.87
AMZN	Amazon	2023-10-23 10:00:00	125.76
AMZN	Amazon	2023-10-23 14:00:00	127.62
AMZN	Amazon	2023-10-23 18:00:00	126.93
AMZN	Amazon	2023-10-24 10:00:00	128.35
AMZN	Amazon	2023-10-24 14:00:00	127.77
AMZN	Amazon	2023-10-24 18:00:00	126.8
AMZN	Amazon	2023-10-25 10:00:00	123.22
AMZN	Amazon	2023-10-25 14:00:00	121.58
AMZN	Amazon	2023-10-25 18:00:00	119.85
AMZN	Amazon	2023-10-26 10:00:00	160.76
AMZN	Amazon	2023-10-26 14:00:00	120.46
AMZN	Amazon	2023-10-26 18:00:00	125.38
AMZN	Amazon	2023-10-27 10:00:00	129.46
AMZN	Amazon	2023-10-27 14:00:00	127.99

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

AMZN	Amazon	2023-10-27 18:00:00	127.76
AMZN	Amazon	2023-10-30 10:00:00	130.63
AMZN	Amazon	2023-10-30 14:00:00	132.64
AMZN	Amazon	2023-10-30 18:00:00	132.43
GOOGL	Google	2023-10-23 10:00:00	136.23
GOOGL	Google	2023-10-23 14:00:00	137.26
GOOGL	Google	2023-10-23 18:00:00	137.23
GOOGL	Google	2023-10-24 10:00:00	139.03
GOOGL	Google	2023-10-24 14:00:00	138.68
GOOGL	Google	2023-10-24 18:00:00	129.99
GOOGL	Google	2023-10-25 10:00:00	126.42
GOOGL	Google	2023-10-25 14:00:00	125.7
GOOGL	Google	2023-10-25 18:00:00	122.63
GOOGL	Google	2023-10-26 10:00:00	123.49
GOOGL	Google	2023-10-26 14:00:00	122.98
GOOGL	Google	2023-10-26 18:00:00	128.15
GOOGL	Google	2023-10-27 10:00:00	121.22
GOOGL	Google	2023-10-27 14:00:00	121.77
GOOGL	Google	2023-10-27 18:00:00	122.27
GOOGL	Google	2023-10-30 10:00:00	124.37
GOOGL	Google	2023-10-30 14:00:00	124.83
GOOGL	Google	2023-10-30 18:00:00	124.55
META	Metaverse	2023-10-23 10:00:00	311.22
META	Metaverse	2023-10-23 14:00:00	315.5
META	Metaverse	2023-10-23 18:00:00	315.24

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

META	Metaverse	2023-10-24 10:00:00	317.36
META	Metaverse	2023-10-24 14:00:00	312.75
META	Metaverse	2023-10-24 18:00:00	310.4
META	Metaverse	2023-10-25 10:00:00	306.48
META	Metaverse	2023-10-25 14:00:00	299.67
META	Metaverse	2023-10-25 18:00:00	290.6
META	Metaverse	2023-10-26 10:00:00	289.65
META	Metaverse	2023-10-26 14:00:00	291.04
META	Metaverse	2023-10-26 18:00:00	290.68
META	Metaverse	2023-10-27 10:00:00	295.95
META	Metaverse	2023-10-27 14:00:00	295.22
META	Metaverse	2023-10-27 18:00:00	296.6
META	Metaverse	2023-10-30 10:00:00	302.73
META	Metaverse	2023-10-30 14:00:00	303.59
META	Metaverse	2023-10-30 18:00:00	302.6
MSFT	Microsoft	2023-10-23 10:00:00	328.28
MSFT	Microsoft	2023-10-23 14:00:00	331.1
MSFT	Microsoft	2023-10-23 18:00:00	330.6
MSFT	Microsoft	2023-10-24 10:00:00	330.78
MSFT	Microsoft	2023-10-24 14:00:00	330.75
MSFT	Microsoft	2023-10-24 18:00:00	343.23
MSFT	Microsoft	2023-10-25 10:00:00	339.54
MSFT	Microsoft	2023-10-25 14:00:00	339.94
MSFT	Microsoft	2023-10-25 18:00:00	338.22
MSFT	Microsoft	2023-10-26 10:00:00	334.08

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

MSFT	Microsoft	2023-10-26 14:00:00	331.46
MSFT	Microsoft	2023-10-26 18:00:00	331.18
MSFT	Microsoft	2023-10-27 10:00:00	334.14
MSFT	Microsoft	2023-10-27 14:00:00	328.9
MSFT	Microsoft	2023-10-27 18:00:00	329.9
MSFT	Microsoft	2023-10-30 10:00:00	334.78
MSFT	Microsoft	2023-10-30 14:00:00	337.06
MSFT	Microsoft	2023-10-30 18:00:00	337.52
TSLA	Tesla	2023-10-23 10:00:00	213.45
TSLA	Tesla	2023-10-23 14:00:00	213.66
TSLA	Tesla	2023-10-23 18:00:00	212.8
TSLA	Tesla	2023-10-24 10:00:00	217.86
TSLA	Tesla	2023-10-24 14:00:00	216.04
TSLA	Tesla	2023-10-24 18:00:00	216.08
TSLA	Tesla	2023-10-25 10:00:00	217.39
TSLA	Tesla	2023-10-25 14:00:00	214.29
TSLA	Tesla	2023-10-25 18:00:00	208.83
TSLA	Tesla	2023-10-26 10:00:00	211.55
TSLA	Tesla	2023-10-26 14:00:00	209.99
TSLA	Tesla	2023-10-26 18:00:00	207.6
TSLA	Tesla	2023-10-27 10:00:00	208.17
TSLA	Tesla	2023-10-27 14:00:00	206.51
TSLA	Tesla	2023-10-27 18:00:00	207.65
TSLA	Tesla	2023-10-30 10:00:00	199.44
TSLA	Tesla	2023-10-30 14:00:00	196.38

TSLA	Tesla	2023-10-30 18:00:00	196.29

3. Which of the rows have a price greater than the average of all prices in the dataset?

The code :

```
WITH aggregates_temp AS (
        SELECT name,
                ROUND(AVG(price), 2) AS avgp, --Assuming 2 decimal for rounding
                MAX(price) AS maxp,
                MIN(price) AS minp,
                (MAX(price) - MIN(price)) AS spread
        FROM stocks
        GROUP BY name
        )
SELECT
        s.name,
        s.symbol,
        s.datetime,
        s.price
FROM stocks s
JOIN aggregates_temp at ON s.name = at.name
WHERE s.price > at.avgp;
```

The result set :

name	symbol	datetime	price
Apple	**AAPL**	**2023-10-23 14:00:00**	**173.5**
Apple	**AAPL**	**2023-10-23 18:00:00**	**172.86**
Apple	**AAPL**	**2023-10-23 10:00:00**	**172.21**
Apple	**AAPL**	**2023-10-24 10:00:00**	**172.53**
Apple	**AAPL**	**2023-10-24 14:00:00**	**173.12**
Apple	**AAPL**	**2023-10-24 18:00:00**	**173.18**
Apple	**AAPL**	**2023-10-25 10:00:00**	**172.16**

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

Apple	**AAPL**	**2023-10-25 14:00:00**	**171.1**
Apple	**AAPL**	**2023-10-30 14:00:00**	**170.5**
Amazon	**AMZN**	**2023-10-26 10:00:00**	**160.76**
Amazon	**AMZN**	**2023-10-27 10:00:00**	**129.46**
Amazon	**AMZN**	**2023-10-30 10:00:00**	**130.63**
Amazon	**AMZN**	**2023-10-30 14:00:00**	**132.64**
Amazon	**AMZN**	**2023-10-30 18:00:00**	**132.43**
Google	**GOOGL**	**2023-10-23 10:00:00**	**136.23**
Google	**GOOGL**	**2023-10-23 14:00:00**	**137.26**
Google	**GOOGL**	**2023-10-23 18:00:00**	**137.23**
Google	**GOOGL**	**2023-10-24 10:00:00**	**139.03**
Google	**GOOGL**	**2023-10-24 14:00:00**	**138.68**
Google	**GOOGL**	**2023-10-24 18:00:00**	**129.99**
Metaverse	**META**	**2023-10-23 10:00:00**	**311.22**
Metaverse	**META**	**2023-10-23 14:00:00**	**315.5**
Metaverse	**META**	**2023-10-23 18:00:00**	**315.24**
Metaverse	**META**	**2023-10-24 10:00:00**	**317.36**
Metaverse	**META**	**2023-10-24 14:00:00**	**312.75**
Metaverse	**META**	**2023-10-24 18:00:00**	**310.4**
Metaverse	**META**	**2023-10-25 10:00:00**	**306.48**
Metaverse	**META**	**2023-10-30 10:00:00**	**302.73**
Metaverse	**META**	**2023-10-30 14:00:00**	**303.59**
Microsoft	**MSFT**	**2023-10-24 18:00:00**	**343.23**
Microsoft	**MSFT**	**2023-10-25 10:00:00**	**339.54**
Microsoft	**MSFT**	**2023-10-25 14:00:00**	**339.94**

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

Microsoft	**MSFT**	**2023-10-25 18:00:00**	**338.22**
Microsoft	**MSFT**	**2023-10-26 10:00:00**	**334.08**
Microsoft	**MSFT**	**2023-10-27 10:00:00**	**334.14**
Microsoft	**MSFT**	**2023-10-30 10:00:00**	**334.78**
Microsoft	**MSFT**	**2023-10-30 14:00:00**	**337.06**
Microsoft	**MSFT**	**2023-10-30 18:00:00**	**337.52**
Tesla	**TSLA**	**2023-10-23 10:00:00**	**213.45**
Tesla	**TSLA**	**2023-10-23 14:00:00**	**213.66**
Tesla	**TSLA**	**2023-10-23 18:00:00**	**212.8**
Tesla	**TSLA**	**2023-10-24 10:00:00**	**217.86**
Tesla	**TSLA**	**2023-10-24 14:00:00**	**216.04**
Tesla	**TSLA**	**2023-10-24 18:00:00**	**216.08**
Tesla	**TSLA**	**2023-10-25 10:00:00**	**217.39**
Tesla	**TSLA**	**2023-10-25 14:00:00**	**214.29**
Tesla	**TSLA**	**2023-10-26 10:00:00**	**211.55**
Tesla	**TSLA**	**2023-10-26 14:00:00**	**209.99**

- ➔ There is a clear pattern of price increasing over the week for the stocks AAPL, AMZN, and GOOGL. The minimum price for each of these stocks was higher on Friday than it was on Monday.

- ➔ There is no clear pattern of price increasing or decreasing over the week for the stocks META, MSFT, and TSLA. The minimum price for each of these stocks was either higher or lower on Friday than it was on Monday.

However, It is important to note that this is just a small sample of data, and it is possible that there would be a different result if we looked at a larger dataset. Additionally, there are many other factors that can affect stock prices, such as economic news, company earnings reports, and investor sentiment.

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

4 .grouping data by the time of the day using strftime() function:

Code :

```
SELECT
  name,
  symbol,
  strftime('%H', datetime) AS hour,
  price
FROM stocks
GROUP BY name, hour, price;
```

Result set :

name	symbol	hour	price
Amazon	AMZN	10	123.22
Amazon	AMZN	10	125.76
Amazon	AMZN	10	128.35
Amazon	AMZN	10	129.46
Amazon	AMZN	10	130.63
Amazon	AMZN	10	160.76
Amazon	AMZN	14	120.46
Amazon	AMZN	14	121.58
Amazon	AMZN	14	127.62
Amazon	AMZN	14	127.77
Amazon	AMZN	14	127.99
Amazon	AMZN	14	132.64
Amazon	AMZN	18	119.85
Amazon	AMZN	18	125.38

Amazon	AMZN	18	126.8
Amazon	AMZN	18	126.93
Amazon	AMZN	18	127.76
Amazon	AMZN	18	132.43
Apple	AAPL	10	168.32
Apple	AAPL	10	169.44
Apple	AAPL	10	172.16
Apple	AAPL	10	172.21
Apple	AAPL	10	172.53
Apple	AAPL	14	167.62
Apple	AAPL	14	167.69
Apple	AAPL	14	170.5
Apple	AAPL	14	171.1
Apple	AAPL	14	173.12
Apple	AAPL	14	173.5
Apple	AAPL	18	164.87
Apple	AAPL	18	166.47
Apple	AAPL	18	167.99
Apple	AAPL	18	169.77
Apple	AAPL	18	172.86
Apple	AAPL	18	173.18
Google	GOOGL	10	121.22
Google	GOOGL	10	123.49
Google	GOOGL	10	124.37
Google	GOOGL	10	126.42

Google	GOOGL	10	136.23
Google	GOOGL	10	139.03
Google	GOOGL	14	121.77
Google	GOOGL	14	122.98
Google	GOOGL	14	124.83
Google	GOOGL	14	125.7
Google	GOOGL	14	137.26
Google	GOOGL	14	138.68
Google	GOOGL	18	122.27
Google	GOOGL	18	122.63
Google	GOOGL	18	124.55
Google	GOOGL	18	128.15
Google	GOOGL	18	129.99
Google	GOOGL	18	137.23
Metaverse	META	10	289.65
Metaverse	META	10	295.95
Metaverse	META	10	302.73
Metaverse	META	10	306.48
Metaverse	META	10	311.22
Metaverse	META	10	317.36
Metaverse	META	14	291.04
Metaverse	META	14	295.22
Metaverse	META	14	299.67
Metaverse	META	14	303.59
Metaverse	META	14	312.75

Metaverse	META	14	315.5
Metaverse	META	18	290.6
Metaverse	META	18	290.68
Metaverse	META	18	296.6
Metaverse	META	18	302.6
Metaverse	META	18	310.4
Metaverse	META	18	315.24
Microsoft	MSFT	10	328.28
Microsoft	MSFT	10	330.78
Microsoft	MSFT	10	334.08
Microsoft	MSFT	10	334.14
Microsoft	MSFT	10	334.78
Microsoft	MSFT	10	339.54
Microsoft	MSFT	14	328.9
Microsoft	MSFT	14	330.75
Microsoft	MSFT	14	331.1
Microsoft	MSFT	14	331.46
Microsoft	MSFT	14	337.06
Microsoft	MSFT	14	339.94
Microsoft	MSFT	18	329.9
Microsoft	MSFT	18	330.6
Microsoft	MSFT	18	331.18
Microsoft	MSFT	18	337.52
Microsoft	MSFT	18	338.22
Microsoft	MSFT	18	343.23

Tesla	TSLA	10	199.44
Tesla	TSLA	10	208.17
Tesla	TSLA	10	211.55
Tesla	TSLA	10	213.45
Tesla	TSLA	10	217.39
Tesla	TSLA	10	217.86
Tesla	TSLA	14	196.38
Tesla	TSLA	14	206.51
Tesla	TSLA	14	209.99
Tesla	TSLA	14	213.66
Tesla	TSLA	14	214.29
Tesla	TSLA	14	216.04
Tesla	TSLA	18	196.29
Tesla	TSLA	18	207.6
Tesla	TSLA	18	207.65
Tesla	TSLA	18	208.83
Tesla	TSLA	18	212.8
Tesla	TSLA	18	216.08

to gain more readability, it's better to use aggregate fonction like AVG().
here is **the code :**

```
SELECT
  name,
  symbol,
  strftime('%H', datetime) AS hour,
  ROUND(AVG(price), 2)  AS avgp
FROM stocks
GROUP BY name, symbol, hour
```

ORDER BY name, hour;

Result set :

name	symbol	hour	avgp
Amazon	AMZN	10	133.03
Amazon	AMZN	14	126.34
Amazon	AMZN	18	126.53
Apple	AAPL	10	170.5
Apple	AAPL	14	170.59
Apple	AAPL	18	169.19
Google	GOOGL	10	128.46
Google	GOOGL	14	128.54
Google	GOOGL	18	127.47
Metaverse	META	10	303.9
Metaverse	META	14	302.96
Metaverse	META	18	301.02
Microsoft	MSFT	10	333.6
Microsoft	MSFT	14	333.2
Microsoft	MSFT	18	335.11
Tesla	TSLA	10	211.31
Tesla	TSLA	14	209.48
Tesla	TSLA	18	208.21

The average price is calculated separately for each stock (name/symbol) at each hour.

Looking at the aggregated results, some noticeable patterns:

- ➔ Amazon (AMZN) and Metaverse (META): Highest prices at 10am, lowest at 6pm
- ➔ Apple (AAPL): Steadily declining price as day goes on
- ➔ Google (GOOGL): Highest at 10am, lowest at 6pm
- ➔ Microsoft (MSFT): Relatively steady across hours
- ➔ Tesla (TSLA): Highest prices at 10am-2pm, lower in morning and evening

So for most of the six stocks, the 10am hour stands out as having the highest average prices, with a gradual decline towards the end of the trading day at 6pm. Microsoft is an exception with a flatter trend.

IV. Using Advanced Queries - - - -

1. Ways to calculate other key statistics about the data, such as the median or variance.

a. First way to calculate median in SQLite using ROW_NUMBER() function:

Code :

```
SELECT
   name,
   symbol,
   AVG(price) AS median
FROM (
   SELECT
      name,
      symbol,
      price,
      ROW_NUMBER() OVER (PARTITION BY name ORDER BY price) AS row_num,
      COUNT(*) OVER (PARTITION BY name) AS total_rows
   FROM stocks
) subquery
WHERE row_num IN ((total_rows + 1) / 2, (total_rows + 2) / 2)
GROUP BY name, symbol;
```

b. Second way to calculate median in SQLite using ROW_NUMBER() function:

The Code :

```
WITH row_table AS (
  SELECT
    name,
    symbol,
    price,
    ROW_NUMBER() OVER (PARTITION BY name ORDER BY price) AS row_num,
    COUNT(*) OVER (PARTITION BY name) AS total_rows
  FROM stocks
)
SELECT
  rt.name,
  rt.symbol,
  ROUND(AVG(CASE
     WHEN row_num IN ((total_rows + 1) / 2, (total_rows + 2) / 2)
     THEN price
  END), 2) AS median
FROM row_table rt
GROUP BY rt.name;
```

Result set :

name	symbol	median
Amazon	AMZN	127.69
Apple	AAPL	170.14
Google	GOOGL	125.27
Metaverse	META	302.67
Microsoft	MSFT	332.77
Tesla	TSLA	210.77

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

2. Getting most aggregates in one table

```
WITH aggregates_temp AS (
  SELECT
    name,
    ROUND(AVG(price), 2) AS avgp,
    MAX(price) AS maxp,
          MIN(price) AS minp,
          (MAX(price) - MIN(price)) AS spread
  FROM stocks
  GROUP BY name
), median_table AS (
  SELECT
   name,
   symbol,
   price,
   ROW_NUMBER() OVER (PARTITION BY name ORDER BY price) AS row_num,
   COUNT(*) OVER (PARTITION BY name) AS total_rows
  FROM stocks
)
SELECT
  at.name,
  at.avgp,
  at.maxp,
  at.minp,
  at.spread,
  ROUND(AVG((price - avgp) * (price - avgp)), 2) AS variance,
  ROUND(AVG(CASE
    WHEN row_num IN ((total_rows + 1) / 2, (total_rows + 2) / 2)
    THEN price
  END), 2) AS median
FROM aggregates_temp at
JOIN median_table mt ON at.name = mt.name
GROUP BY at.name;
```

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

Result set :

name	avgp	maxp	minp	spread	variance	median
Amazon	128.63	160.76	119.85	40.91	73.24	127.69
Apple	170.09	173.5	164.87	8.63	6.52	170.14
Google	128.16	139.03	121.22	17.81	39.6	125.27
Metaverse	302.63	317.36	289.65	27.71	84.94	302.67
Microsoft	333.97	343.23	328.28	14.95	17.9	332.77
Tesla	209.67	217.86	196.29	21.57	41.94	210.77

Analyzing aggregates

Amazon (AMZN):
- High variance at 73.24 indicating volatile prices
- Large spread between min and max prices
- Median close to average showing normal distribution

Apple (AAPL):
- Very low variance
- Narrow spread
- Median and average approximately equal
  ➔ This shows stable prices with minimal volatility

Google (GOOGL):
- Higher variance at 39.6
- Median lower than average showing negative skew
- Wider spread of prices
  ➔ More volatile, skewed by extremely high prices

Metaverse (META):
- Similar shape to Google
- High variance
- Lower median than average
- Wider price spread
  ➔ Negatively skewed distribution and high volatility

Data Science Independent Project #1 – Watching the Stock Market

Performed by Abdessamad HAJIBI, Codecademy, Career Path : Data Scientist : Analytics Specialist

Microsoft (MSFT):
- Low variance
- Very narrow spread
- Average and median aligned
    ➔ Stable narrow range of prices

Tesla (TSLA):
- High variance
- Wider price spread
    ➔ Considerable volatility

In summary - the variance, spread and median vs. average comparisons reveal insights into price stability and distribution shapes for each stock over the trading period. Please let me know if this analysis properly covers the updated table with variance included!