**COCACOLA STOCK ANALYSIS**

**PROJECT-2 SQL QUERIES**

Query 1

copy cocacola\_stock

from 'G:/SQL/project\_cococola/cocacola\_data.csv'

delimiter ','

csv header;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

select \* from cocacola\_stock

limit 10

Query 2

alter table cocacola\_stock

alter column "Date" type date

using "Date"::date,

alter column "Open" type numeric(10,2)

using "Open"::numeric,

alter column "High" type numeric(10,2)

using "High"::numeric,

alter column "Low" type numeric(10,2)

using "Low"::numeric,

alter column "Close" type numeric(10,2)

using "Close"::numeric,

alter column "Adj Close" type numeric(10,2)

using "Adj Close"::numeric,

alter column "Volume" type Bigint

using "Volume"::Bigint

Query 3 –Cleaning—

select \* from cocacola\_stock

where 'Date' is null

or 'Open' is null

or 'High' is null

or 'Low' is null

or 'Close' is null

or 'Adj Close' is null

or 'Volume' is null

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

select \* from cocacola\_stock

where 'Date'='0'

or 'Open' = '0'

or 'High' = '0'

or 'Low' = '0'

or 'Close' = '0'

or 'Adj Close'= '0'

or 'Volume'= '0'

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

select \* from cocacola\_stock

where "Low"> "High"

--duplicates—

select "Date", count(\*) from cocacola\_stock

group by "Date"

having count(\*)>1

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

--open and close sanity check--

select \* from cocacola\_stock

where "Open"> "High"

or "Open"< "Low"

or "Close"> "High"

or "Close"< "Low"

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

--check unrealistic spikes--

select \* from cocacola\_stock

where "Volume"> 1000000000

order by "Volume" Desc

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

--check negative values--

select \*

from cocacola\_stock

where "High"<0

or "Low"<0

or "Close"<0

or "Adj Close"<0

or "Volume"<0

Query 4 –data validation—

--diff between close and adjclose---

select \*, ("Adj Close"-"Close")as diff from cocacola\_stock

where "Adj Close"<>"Close"

order by "Date"

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

--identical close and adjclose--

select count(\*) as totalrow, count(\*) filter(where "Close"="Adj Close")as matched\_rows

from cocacola\_stock

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

--cronological order check--

select \* from cocacola\_stock

order by "Date" asc

--Query5-- solving business insights questions--

--data analysis--

--1. Trend Over Time--

select "Date", "Adj Close"

from cocacola\_stock

order by "Date"

--2. Highest and Lowest Prices--

select Max("High")as Highest, Min("Low")as Lowest

from cocacola\_stock

--3. On which day did Coca-Cola stock gain the most, and on which day did it fall the most? (Daily return %).—--value-pre.value/pre.value\*100--

--Gain—

Select "Date", "Adj Close",

round((("Adj Close"-lag("Adj Close") over(order by "Date"))/lag("Adj Close") over(order by "Date"))\*100,2) as gain

from cocacola\_stock

order by gain desc nulls Last

limit 1

--Fall--

Select "Date", "Adj Close",

round((("Adj Close"-lag("Adj Close") over(order by "Date"))/lag("Adj Close") over(order by "Date"))\*100,2) as gain

from cocacola\_stock

order by gain asc nulls Last

limit 1

4. What is the average closing price per month and per year?

--Month—

select date\_trunc('month', "Date") as mdate, round(avg("Adj Close"),2) as avg

from cocacola\_stock

group by 1

order by 1

--Year—

select date\_trunc('year', "Date") as ydate, round(avg("Adj Close"),2) as avg

from cocacola\_stock

group by 1

order by 1

5. Which days had the highest and lowest trading volumes?, Is there a trend in trading volume over time?

-- Trading Volume Patterns—

select "Date", "Volume"

from cocacola\_stock

order by "Volume" DESC

limit 5

select "Date", "Volume"

from cocacola\_stock

order by "Volume" asc

limit 5

6. Which days had the largest difference between High and Low prices?

--Volatility Check—

select "Date", ("High"-"Low")as Volatility

from cocacola\_stock

order by Volatility Desc

limit 5

7. What is the 7days and 30days moving averages of Adj Close ?

--Moving Averages—

select "Date", "Adj Close",

round(avg("Adj Close")over(order by "Date" rows between 6 preceding and current row),2)as weekly\_avg,

round(avg("Adj Close") over(order by "Date" rows between 29 preceding and current row),2)as monthly\_avg

from cocacola\_stock

order by "Date"

bullish or bearish trends?

select "Date", "Adj Close",

round(avg("Adj Close")over(order by "Date" rows between 6 preceding and current row),2)as weekly\_avg,

round(avg("Adj Close") over(order by "Date" rows between 29 preceding and current row),2)as monthly\_avg,

case

when round(avg("Adj Close")over(order by "Date" rows between 6 preceding and current row),2)

>round(avg("Adj Close") over(order by "Date" rows between 29 preceding and current row),2)

then 'Bullish'

else 'Bearish'

end

from cocacola\_stock

order by "Date"

8. How often does the stock close higher than it opens?

select count(\*) filter(where "Close">"Open")as higher\_close,

count(\*) filter(where "Close"<"Open")as lower\_close,

count(\*) filter(where "Close"="Open")as same

from cocacola\_stock

9. What is the distribution of daily /monthly/quarterly/yearly returns %?

--Returns Daily—

select "Date",Round((("Adj Close"-lag("Adj Close")over(order by "Date"))/lag("Adj Close")over(order by "Date"))

\*100, 2)as daily\_returns

from cocacola\_stock

order by "Date"

--OR--without null

with mycte as(select "Date", "Adj Close", lag("Adj Close")over(order by "Date")as pre\_close

from cocacola\_stock)

select "Date", round((("Adj Close"-pre\_close)/pre\_close)\*100, 2)||'%' as daily\_returns

from mycte

where pre\_close is not null

---Returns Monthly—

--find first and last Adj Close of each month—

select date\_trunc('month',"Date")::date as mdate, --without casting mdate shows timestamp--

round(((max("Adj Close")-min("Adj Close"))/min("Adj Close"))\*100, 2) ||'%' as monthly\_returns

from cocacola\_stock

group by 1

order by 2 desc

---Returns Quarterly—

--find first and last Adj Close of each quarter--

select date\_trunc('quarter',"Date")::date as qdate, --without casting date it shows timestamp--

round(((max("Adj Close")-min("Adj Close"))/min("Adj Close"))\*100, 2) ||'%' as quarterly\_returns

from cocacola\_stock

group by 1

order by 2 desc

---Returns Yearly—

--find first and last Adj Close of each Year--

select date\_trunc('year',"Date")::date as ydate, --without casting date it shows timestamp--

round(((max("Adj Close")-min("Adj Close"))/min("Adj Close"))\*100, 2)||'%' as yearly\_returns

from cocacola\_stock

group by 1

order by 1

10. Does Coca-Cola stock show any seasonal patterns /stronger performance in certain months/quarters?

--Seasonal / Periodic Analysis—

--month--

select extract(month from "Date")as mdate, round(avg("Adj Close"),2)as avg\_adj

from cocacola\_stock

group by 1

order by 2 desc

--quarter--

select extract(quarter from "Date")as qdate, round(avg("Adj Close"),2)as avg\_adj

from cocacola\_stock

group by 1

order by 2 desc