

Assignment E9

Name Praveen

Roll no 22b3931

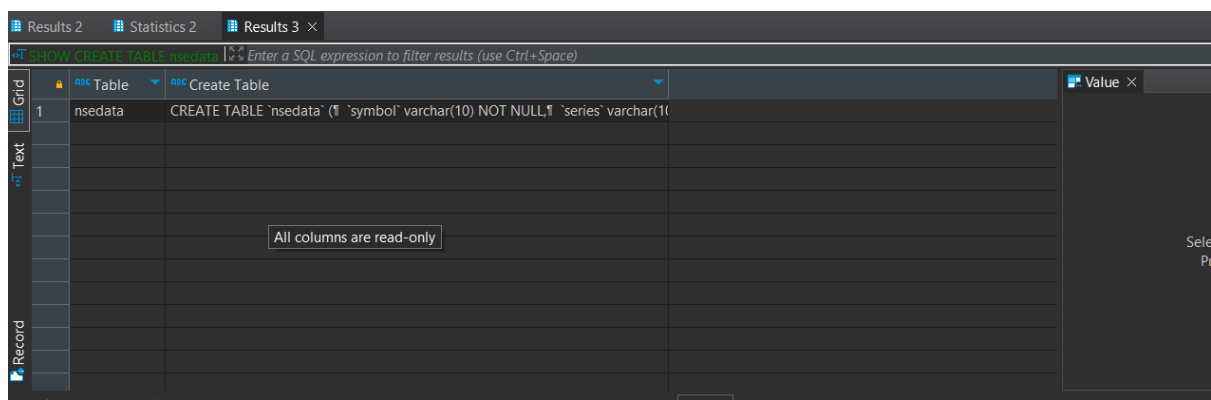
Branch electrical dual

2. Select the database stock data using SQL

```
USE stockdata;
```

3. Get a schema dump of the table nsedata

```
SHOW CREATE TABLE nsedata;
```

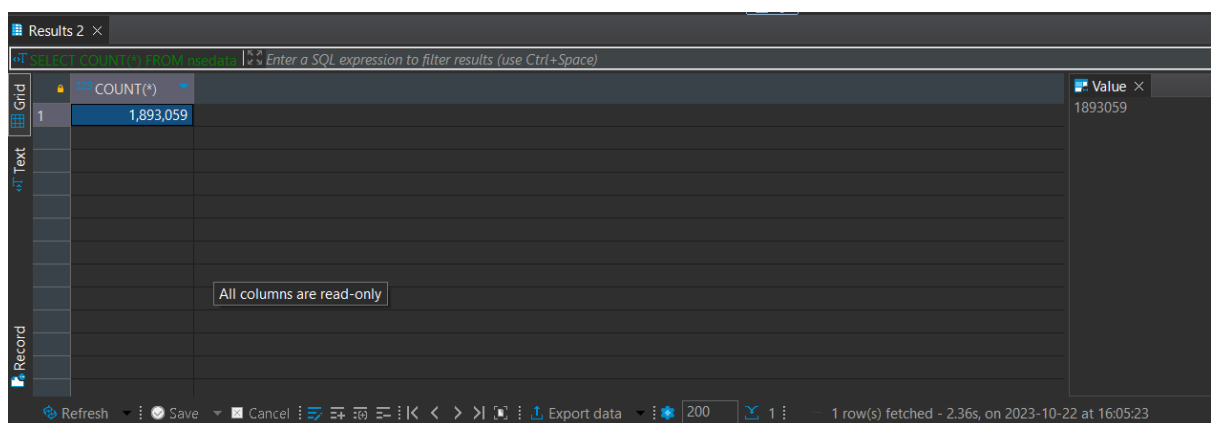


The screenshot shows a database client interface with a results grid. The grid has three columns: 'Table', 'Create Table', and 'Value'. The first row shows the table 'nsedata' and its creation statement: 'CREATE TABLE `nsedata` (`symbol` varchar(10) NOT NULL, `series` varchar(10) NOT NULL, `timestamp` timestamp(6) NOT NULL, `open` float(12) NOT NULL, `high` float(12) NOT NULL, `low` float(12) NOT NULL, `close` float(12) NOT NULL, `volume` int(11) NOT NULL) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci'. A message 'All columns are read-only' is displayed in the center of the grid. The 'Value' column is empty.

Table	Create Table	Value
nsedata	CREATE TABLE `nsedata` (`symbol` varchar(10) NOT NULL, `series` varchar(10) NOT NULL, `timestamp` timestamp(6) NOT NULL, `open` float(12) NOT NULL, `high` float(12) NOT NULL, `low` float(12) NOT NULL, `close` float(12) NOT NULL, `volume` int(11) NOT NULL) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci	

4. Get a count of the total number of records in nsedata

```
SELECT COUNT(*) FROM nsedata;
```



The screenshot shows a database client interface with a results grid. The grid has two columns: 'COUNT(*)' and 'Value'. The first row shows the count '1,893,059'. A message 'All columns are read-only' is displayed in the center of the grid. The 'Value' column shows the value '1893059'.

COUNT(*)	Value
1,893,059	1893059

5. Get the total count of the records for the month “October 2012”

```
SELECT COUNT(*) FROM nsedata WHERE timestamp = '01-OCT-2012';
```

Results 2 Statistics 2 Results 3 ×

Enter a SQL expression to filter results (use Ctrl+Space)

Grid

1 1,585

Value × 1585

Refresh Save Cancel Export data 200 1 1 row(s) fetched - 4.167s, on 2023-10-22 at 17:47:03

6. Repeat '4', but only for the stock with symbol "GEOMETRIC"

```
SELECT COUNT(*) FROM nsedata WHERE symbol = 'GEOMETRIC';
```

Results 2 Statistics 2 Results 3 ×

Enter a SQL expression to filter results (use Ctrl+Space)

Grid

1 1,237

Value × 1237

Refresh Save Cancel Export data 200 1 1 row(s) fetched - 1.741s, on 2023-10-22 at 17:51:31

7. Repeat '6', but only display the first 10 records

```
SELECT * FROM nsedata WHERE symbol = 'GEOMETRIC' LIMIT 10;
```

Results 2 Statistics 2 nsedata 3 ×

Enter a SQL expression to filter results (use Ctrl+Space)

	symbol	series	123 open	123 high	123 low	123 close	123 last	123 prevclose	123 tottrdqty	123 tottrdval
1	GEOMETRIC	EQ	62.35	64.5	61.4	63.25	63.25	61.3	82,246	5,179,345.65
2	GEOMETRIC	EQ	100.7	105.5	99.1	103.5	102.55	100.2	124,482	12,753,266.8
3	GEOMETRIC	EQ	116	121	116	120	120.2	115.55	644,060	77,015,430
4	GEOMETRIC	EQ	166.5	184.5	163	177.55	177.4	167.15	2,398,121	426,671,089.7
5	GEOMETRIC	EQ	49.8	50.1	49.05	49.9	50	48.9	55,376	2,765,041.05
6	GEOMETRIC	EQ	94.4	94.65	90.4	91.8	91.95	94.2	887,542	82,336,516.15
7	GEOMETRIC	EQ	69.45	70.05	63	64.1	63.3	69.45	319,336	20,718,691.55
8	GEOMETRIC	EQ	141.2	144	139.5	140.2	140.5	142.6	291,911	41,292,337.35
9	GEOMETRIC	EQ	73.3	73.6	71.35	72.35	72.4	72.15	130,567	9,479,917.2
10	GEOMETRIC	EQ	45.9	48.9	45.5	47.2	47	45	124,440	5,906,044.7

Refresh Save Cancel Export data 200 10 10 row(s) fetched - 547ms, on 2023-10-22 at 17:51:31

8. Totally, how many records of "INFY" does the table contain?

```
SELECT COUNT(*) FROM nsedata WHERE symbol = 'INFY';
```

Results 2 Statistics 2 Results 3 ×

SELECT COUNT(*) FROM rsedata WHERE symbol = 'INFX' Enter a SQL expression to filter results (use Ctrl+Space)

Grid

	COUNT(*)
1	1,023

Text

Record

Value ×

1023

Refresh Save Cancel

Export data 200 1

1 row(s) fetched - 1.638s, on 2023-10-22 at 17:59:51

9. Get a listing of the first 10 records of “3IINFOTECH”, but the listing should contain only the following columns: symbol, open, high, low, close, and timestamp

```
SELECT symbol, open, high, low, close, timestamp FROM nsedata
WHERE symbol = '3IINFOTECH'
LIMIT 10;
```

Results 2Statistics 2nsedata 3

SQL

SELECT symbol, open, high, low, close, timestamp FROM ;

Enter a SQL expression to filter results (use Ctrl+Space)

Grid

Text

Record

	symbol	open	high	low	close	timestamp	
1	3IINFOTECH	43.75	45.3	43.75	44.9	01-APR-2011	
2	3IINFOTECH	5.65	6.1	5.65	6.1	01-APR-2013	
3	3IINFOTECH	7.85	7.9	7.45	7.65	01-APR-2014	
4	3IINFOTECH	5.9	6.3	5.8	6.2	01-APR-2015	
5	3IINFOTECH	41.6	42.45	40.2	40.45	01-AUG-2011	
6	3IINFOTECH	10.8	10.8	10.5	10.8	01-AUG-2012	
7	3IINFOTECH	3.95	4.15	3.85	4	01-AUG-2013	
8	3IINFOTECH	8.75	9.1	8.6	8.65	01-AUG-2014	
9	3IINFOTECH	55.9	59.4	55.55	58.35	01-DEC-2010	
10	3IINFOTECH	20	20	18.5	18.65	01-DEC-2011	

10. Repeat '9', but this time use the results to create a temporary table t1

```
CREATE TEMPORARY TABLE t1 AS
SELECT symbol, open, high, low, close, timestamp
FROM nsedata
WHERE symbol = '3IINFOTECH'
LIMIT 10;
SELECT * FROM t1;
```

Results 2 Statistics 2 t1 3 ×							
SELECT * FROM t1 Enter a SQL expression to filter results (use Ctrl+Space)							
	symbol	open	high	low	close	timestamp	
1	3IINFOTECH	43.75	45.3	43.75	44.9	01-APR-2011	
2	3IINFOTECH	5.65	6.1	5.65	6.1	01-APR-2013	
3	3IINFOTECH	7.85	7.9	7.45	7.65	01-APR-2014	
4	3IINFOTECH	5.9	6.3	5.8	6.2	01-APR-2015	
5	3IINFOTECH	41.6	42.45	40.2	40.45	01-AUG-2011	
6	3IINFOTECH	10.8	10.8	10.5	10.8	01-AUG-2012	
7	3IINFOTECH	3.95	4.15	3.85	4	01-AUG-2013	
8	3IINFOTECH	8.75	9.1	8.6	8.65	01-AUG-2014	
9	3IINFOTECH	55.9	59.4	55.55	58.35	01-DEC-2010	
10	3IINFOTECH	20	20	18.5	18.65	01-DEC-2011	

11. Using t1 find out the following for the column close: max, min, mean, standard deviation and variance

```
SELECT
  MAX(close) AS max_close,
  MIN(close) AS min_close,
  AVG(close) AS mean_close,
  STDDEV(close) AS std_dev,
  VARIANCE(close) AS variance
FROM t1;
```

Results 1 ×						
SELECT MAX(close) AS max_close, MIN(close) AS min_close, AVG(close) AS mean_close, STDDEV(close) AS std_dev, VARIANCE(close) AS variance FROM t1; Enter a SQL expression to filter results (use Ctrl+Space)						
	max_close	min_close	mean_close	std_dev	variance	
1	58.35	4	20.575	18.7432287773	351.308625	

12. How will you find out the value of the median?

Median can be found using below snippet of code

```
SELECT
  PERCENTILE_CONT(0.5) WITHIN GROUP (ORDER BY close) AS median
FROM t1;
```

13. Delete table t1

```
DROP TABLE t1;
```

14. Use nsedata. Using the GROUP BY functionality of SQL create a table t2 containing the average value of close for each and every symbol in the table. Hint: the table will have the columns: symbol, average

```
DROP TABLE t1;
CREATE TABLE t2 AS
SELECT symbol, AVG(close) AS average
FROM nsedata
GROUP BY symbol;
```

The screenshot shows a SQL IDE with a query editor, an output pane, and a statistics pane.

Query Editor:

```

LIMIT 10;
SELECT * FROM t1;

SELECT
    MAX(close) AS max_close,
    MIN(close) AS min_close,
    AVG(close) AS mean_close,
    STDDEV(close) AS std_dev,
    VARIANCE(close) AS variance
FROM t1;

DROP TABLE t1;
CREATE TABLE t2 AS
SELECT symbol, AVG(close) AS average
FROM nsedata
GROUP BY symbol;

```

Output Pane:

Enter a part of a message to search for here

Data truncated for column 'average' at row 1117
 Data truncated for column 'average' at row 1118
 Data truncated for column 'average' at row 1119
 Data truncated for column 'average' at row 1120
 Data truncated for column 'average' at row 1121
 Data truncated for column 'average' at row 1122
 Data truncated for column 'average' at row 1123
 Data truncated for column 'average' at row 1124
 Data truncated for column 'average' at row 1125
 Data truncated for column 'average' at row 1126
 Data truncated for column 'average' at row 1127
 Data truncated for column 'average' at row 1128
 Data truncated for column 'average' at row 1130
 Data truncated for column 'average' at row 1131
 Data truncated for column 'average' at row 1132
 Data truncated for column 'average' at row 1133

Statistics 1

Name	Value
Updated Rows	2048
Query	CREATE TABLE t2 AS SELECT symbol, AVG(close) AS average FROM nsedata GROUP BY symbol
Start time	Sun Oct 22 23:02:51 IST 2023
Finish time	Sun Oct 22 23:03:09 IST 2023

15. Create a table t3 such that it contains the following columns: symbol, open, close, "average of open and close". Fill up this table for the company GEOMETRIC, for the month of October 2012

```
CREATE TABLE t3 AS
SELECT
    symbol,
    open,
    close,
    (open + close) / 2 AS average_open_close
FROM
    nsedata
WHERE
    symbol = 'GEOMETRIC'
    AND STR_TO_DATE(timestamp, '%d-%M-%Y') BETWEEN '2012-10-01' AND '2012-10-31';
```

16. It is required to create a table t4 such that it contains the data for two companies GEOMETRIC and TCS. The columns of this table should be as follows: timestamp, close_tcs, close_geometric. Hint: use JOIN

```
CREATE TABLE t4 AS
SELECT
    A.timestamp,
    A.close AS close_geometric,
    B.close AS close_tcs
FROM
```

```

nsedata A
JOIN
nsedata B
ON
A.timestamp = B.timestamp
WHERE
A.symbol = 'GEOMETRIC'
AND B.symbol = 'TCS';

```

17. Find out the maximum and minimum difference in the daily closing prices of these two companies.

```
SELECT
    MAX(close_geometric - close_tcs) AS max_difference,
    MIN(close_geometric - close_tcs) AS min_difference
FROM
    t4;
```

Results 1

Enter a SQL expression to filter results (use Ctrl+Space)

	max_difference	min_difference
1	-770.35	-2,631.65

All columns are read-only

Value
-770.350000

18. Based on t_4 can you identify those days on which the difference in their closing price was more than the average of the minimum and maximum difference.

```
SELECT
    timestamp
FROM
    t4
WHERE
    close_geometric - close_tcs > (
        (SELECT MAX(close_geometric - close_tcs) FROM t4) +
        (SELECT MIN(close_geometric - close_tcs) FROM t4)
    ) / 2;
```

t4 1 x	
SELECT timestamp FROM t4 WHERE close_geometric < 200	
Grid	timestamp
1	01-APR-2011
2	01-APR-2013
3	01-AUG-2011
4	01-AUG-2012
5	01-DEC-2010
6	01-DEC-2011
7	01-FEB-2011
8	01-FEB-2012
9	01-FEB-2013
10	01-JAN-2013
11	01-JUL-2011
12	01-JUL-2013
13	01-JUN-2011

19. Based on nsedata, create table t5 such that it contains the average close price of each company traded in the month of April 2012. The table should be sorted in descending order of the average close price.

```
CREATE TABLE t5 AS
SELECT
    symbol,
    open,
    close,
    (open + close) / 2 AS average_open_close
FROM
    nsedata
WHERE
    symbol = 'GEOMETRIC'
    AND DATE_FORMAT(STR_TO_DATE(timestamp, '%d-%M-%Y'), '%Y-%m-%d') BETWEEN
    '2012-10-01' AND '2012-10-31';
```

20. Not all companies are traded every day. It is required to create a table that contains a count of the days each company has been traded. The table should be sorted in descending order of the count

```
CREATE TABLE t6 AS
SELECT
    symbol,
    open,
    close,
    (open + close) / 2 AS average_open_close
FROM
    nsedata
WHERE
    symbol = 'GEOMETRIC'
    AND DATE(STR_TO_DATE(timestamp, '%d-%M-%Y')) BETWEEN '2012-10-01' AND
    '2012-10-31';
```