Solving an SDG Problem with Data (Choose Your SDG)

Part 1: SDG Selection and Problem Definition

- **SDG Selection:** Choose an SDG (e.g., SDG 3: Good Health, SDG 7: Affordable and Clean Energy).

Good Health and Well-being (SDG 3)

- **Problem Definition:** Define a specific problem within your chosen SDG that can be addressed using data.

A specific problem within SDG 3: Good Health and Well-being that can be addressed using data is the high rate of missed medical appointments (no-shows) in healthcare facilities. This issue can lead to inefficiencies in resource utilization, increased healthcare costs, and poorer health outcomes. By analyzing data on missed appointments, healthcare providers can identify patterns and factors contributing to no-shows, enabling them to implement targeted interventions to reduce missed appointments and improve patient care.

Part 2: Database Design

- **ERD:** Design an ERD for your project, including entities relevant to your SDG problem.

https://drive.google.com/file/d/1-dNc860-r4ChmBdfPHsiFBPKfrXUHvA4/view?usp=drive_link

- **Schema:** Write SQL statements to create the database schema based on your ERD.

```
CREATE DATABASE noshow;


use noshow;

CREATE TABLE NoShowAppointments (

    PatientId BIGINT,

    AppointmentID BIGINT,

    Gender CHAR(1),

    ScheduledDay DATETIME,
    AppointmentDay DATE,

    Age INT,

    Neighbourhood VARCHAR(100),

    Scholarship VARCHAR(3),

    Hipertension VARCHAR(3),

    Diabetes VARCHAR(3),

    Alcoholism VARCHAR(3),

    Handcap VARCHAR(3),

    SMS_received VARCHAR(3),

    No_show VARCHAR(3)

);
```

- **Sample Data:** Populate the database with relevant sample data.

```
select* from noshowappointments;


INSERT INTO NoShowAppointments (PatientId, AppointmentID, Gender,
ScheduledDay, AppointmentDay, Age, Neighbourhood, Scholarship,
Hipertension, Diabetes, Alcoholism, Handcap, SMS_received, No_show) VALUES

(29872500000000, 5642903, 'F', '2016-04-29 15:10:45', '2016-04-29', 62,
'JARDIM DA PENHA', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No'),

(558998000000000, 5642503, 'M', '2016-04-29 16:08:27', '2016-04-29', 56,
'JARDIM DA PENHA', 'No', 'No', 'No', 'No', 'No', 'No', 'No'),

(4262960000000, 5642549, 'F', '2016-04-29 16:19:04', '2016-04-29', 62,
'MATA DA PRAIA', 'No', 'No', 'No', 'No', 'No', 'No', 'No'),

(867951000000, 5642828, 'F', '2016-04-29 17:29:31', '2016-04-29', 8,
'PONTAL DE CAMBURI', 'No', 'No', 'No', 'No', 'No', 'No', 'No'),

(8841190000000, 5642494, 'F', '2016-04-29 16:07:23', '2016-04-29', 56,
'JARDIM DA PENHA', 'No', 'Yes', 'Yes', 'No', 'No', 'No', 'No'),

(95985100000000, 5626772, 'F', '2016-04-27 08:36:51', '2016-04-29', 76,
'REPÚBLICA', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No'),

(733688000000000, 5630279, 'F', '2016-04-27 15:05:12', '2016-04-29', 23,
'GOIABEIRAS', 'No', 'No', 'No', 'No', 'No', 'No', 'Yes'),

(3449833394123, 5630575, 'F', '2016-04-27 15:39:58', '2016-04-29', 39,
'GOIABEIRAS', 'No', 'No', 'No', 'No', 'No', 'No', 'Yes'),

(56394729949972, 5638447, 'F', '2016-04-29 08:02:16', '2016-04-29', 21,
'ANDORINHAS', 'No', 'No', 'No', 'No', 'No', 'No', 'Yes'),

(78124564369297, 5629123, 'F', '2016-04-27 12:48:25', '2016-04-29', 19,
'CONQUISTA', 'No', 'No', 'No', 'No', 'No', 'No', 'No'),
(734536231958495, 5630213, 'F', '2016-04-27 14:58:11', '2016-04-29', 30,
'NOVA PALESTINA', 'No', 'No', 'No', 'No', 'No', 'No', 'No'),

(7542951368435, 5620163, 'M', '2016-04-26 08:44:12', '2016-04-29', 29,
'NOVA PALESTINA', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(3737429732412, 5574527, 'F', '2016-04-12 15:10:45', '2016-04-29', 34,
'SANTO ANTÔNIO', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(377242196277, 5574534, 'F', '2016-04-12 15:11:24', '2016-04-29', 56,
'SANTO ANTÔNIO', 'No', 'Yes', 'No', 'No', 'No', 'Yes', 'No'),

(9538868114712, 5594665, 'F', '2016-04-18 11:05:04', '2016-04-29', 59,
'SANTO ANTÔNIO', 'No', 'Yes', 'Yes', 'No', 'No', 'Yes', 'No'),

(59998167795865, 5600591, 'M', '2016-04-19 10:55:18', '2016-04-29', 8,
'NOVA PALESTINA', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(35856213466619, 5591384, 'F', '2016-04-15 16:58:02', '2016-04-29', 15,
'BELA VISTA', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(847237964596, 5641363, 'F', '2016-04-29 13:24:28', '2016-04-29', 54,
'SANTO ANTÔNIO', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No'),

(4673335446385, 5574680, 'F', '2016-04-12 15:35:30', '2016-04-29', 18,
'SANTO ANTÔNIO', 'No', 'No', 'No', 'No', 'No', 'Yes', 'Yes'),

(744463183, 5539219, 'F', '2016-04-01 14:27:30', '2016-04-29', 34, 'SANTO
ANTÔNIO', 'Yes', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(9711615822964, 5539858, 'F', '2016-04-01 16:01:28', '2016-04-29', 35,
'SANTO ANTÔNIO', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(329315252114, 5546448, 'F', '2016-04-05 16:14:11', '2016-04-29', 33,
'MÁRIO CYPRESTE', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(6297228525237, 5590755, 'F', '2016-04-15 14:54:29', '2016-04-29', 16,
'MÁRIO CYPRESTE', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(456431233657599, 5540024, 'F', '2016-04-01 16:40:41', '2016-04-29', 18,
'MÁRIO CYPRESTE', 'No', 'No', 'No', 'No', 'No', 'Yes', 'No'),

(4599699374244, 5642643, 'M', '2016-04-29 16:34:59', '2016-04-29', 46,
'BELA VISTA', 'No', 'Yes', 'No', 'No', 'No', 'No', 'No'),

(36732783669262, 5621388, 'F', '2016-04-26 10:12:46', '2016-04-29', 31,
'BELA VISTA', 'Yes', 'No', 'No', 'No', 'No', 'Yes', 'Yes'),

(547545332189, 5642319, 'F', '2016-04-29 15:35:28', '2016-04-29', 7, 'SANTO
ANTÔNIO', 'No', 'No', 'No', 'No', 'No', 'No', 'No'),

(99972128493541, 5542592, 'F', '2016-04-05 09:09:15', '2016-04-29', 36,
'MÁRIO CYPRESTE', 'No', 'Yes', 'No', 'No', 'No', 'Yes', 'No'),

(944758151797235, 5633576, 'F', '2016-04-28 09:37:22', '2016-04-29', 67,
'PRAIA DO SUÁ', 'No', 'No', 'Yes', 'No', 'No', 'No', 'Yes'),

(4224165869815, 5561194, 'M', '2016-04-08 10:14:36', '2016-04-29', 42,
'PRAIA DO SUÁ', 'No', 'Yes', 'Yes', 'No', 'No', 'Yes', 'No'),

(948869664473423, 5571842, 'F', '2016-04-12 09:41:01', '2016-04-29', 34,
'PRAIA DO SUÁ', 'Yes', 'No', 'No', 'No', 'No', 'Yes', 'Yes'),
(55373667525673, 5560735, 'F', '2016-04-08 09:28:50', '2016-04-29', 22,
'PRAIA DO SUÁ', 'No', 'No', 'No', 'No', 'No', 'Yes', 'Yes'),

(152399288918984, 5639855, 'F', '2016-04-29 09:42:22', '2016-04-29', 5,
'SANTA HELENA', 'Yes', 'No', 'No', 'No', 'No', 'No', 'No'),

(47559375448498, 5637150, 'M', '2016-04-28 17:38:06', '2016-04-29', 29,
'PRAIA DO SUÁ', 'No', 'No', 'No', 'Yes', 'No', 'No', 'Yes'),

(846858655833146, 5560742, 'M', '2016-04-08 09:29:23', '2016-04-29', 0,
'PRAIA DO SUÁ', 'No', 'No', 'No', 'No', 'No', 'Yes', 'Yes'),

(3587185996856, 5580520, 'M', '2016-04-14 07:07:10', '2016-04-29', 69,
'PRAIA DO SUÁ', 'No', 'No', 'No', 'No', 'No', 'Yes', 'Yes');
```

Part 3: SQL Programming

- **Data Retrieval:** Write SQL queries to retrieve relevant data based on your problem definition.

1. Retrieve the total number of missed appointments (no-shows):

```
SELECT COUNT(*) AS TotalNoShows

FROM NoShowAppointments

WHERE No_show = 'Yes';
```

2. Retrieve the number of no-shows by patient:

```
SELECT PatientID, COUNT(*) AS NoShowCount

FROM NoShowAppointments

WHERE No_show = 'Yes'

GROUP BY PatientID;
```

3. Retrieve the number of no-shows by appointment day:

```
SELECT AppointmentDay, COUNT(*) AS NoShowCount

FROM NoShowAppointments

WHERE No_show = 'Yes'

GROUP BY AppointmentDay

ORDER BY AppointmentDay;
```

4. Retrieve the average of patient’s age;

```
SELECT AVG(Age) from noshowappointments;
```

5. Retrieve minimum age of patients

```
SELECT MAX(Age) from noshowappointments;
```

6. Retrieve the neighbourhoods

```
SELECT DISTINCT Neighbourhood FROM noshowappointments;
```

- **Data Analysis:** Write SQL queries to analyze data and generate insights related to your SDG problem.

1. Analysis of the no-show rate for different age groups:

```
SELECT
    CASE
        WHEN Age BETWEEN 0 AND 17 THEN '0-17'
        WHEN Age BETWEEN 18 AND 35 THEN '18-35'
        WHEN Age BETWEEN 36 AND 50 THEN '36-50'
        WHEN Age BETWEEN 51 AND 65 THEN '51-65'
        ELSE '66+'
    END AS AgeGroup,
    COUNT(*) AS NoShowCount
FROM NoShowAppointments
WHERE No_show = 'Yes'
GROUP BY AgeGroup
ORDER BY AgeGroup ASC;
```

2. Analysis of number of no-shows by gender:

```
SELECT Gender, COUNT(*) AS NoShowCount
FROM NoShowAppointments
WHERE No_show = 'Yes'
GROUP BY Gender;
```

3. Analysis of Reminder Effectiveness

```
SELECT SMS_received, COUNT(*) AS NoShowCount
FROM NoShowAppointments
WHERE No_show = 'Yes'
```

```
GROUP BY SMS_received;
```

4. Analysis of Neighbourhood Effectiveness

```
SELECT Neighbourhood, COUNT(*) AS NoShowCount
FROM NoShowAppointments
WHERE No_show = 'Yes'
GROUP BY Neighbourhood
ORDER BY NoShowCount DESC;
```

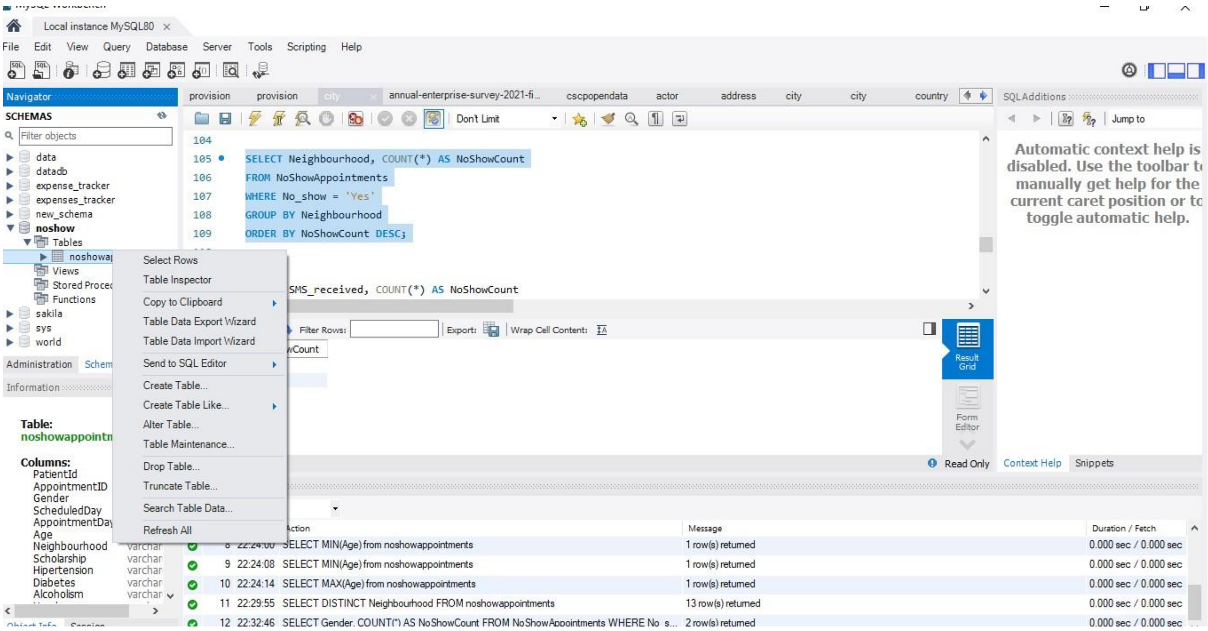
Part 4: Data Analysis Using Excel

- **Import Data:** Import data from your database into Excel.
- **Analysis:** Analyze the data using pivot tables, charts, and other Excel tools.
- **Dashboard:** Create an interactive Excel dashboard to visualize key insights.

Part 5: Integration and Testing

- **Integration:** Document the process of importing data into Excel and ensuring consistency.

Right click the noshowappointments table in MySQL Workbench and click on **Table Data Export Wizard.**



```
SELECT Neighbourhood, COUNT(*) AS NoShowCount
FROM NoShowAppointments
WHERE No_show = 'Yes'
GROUP BY Neighbourhood
ORDER BY NoShowCount DESC;
```

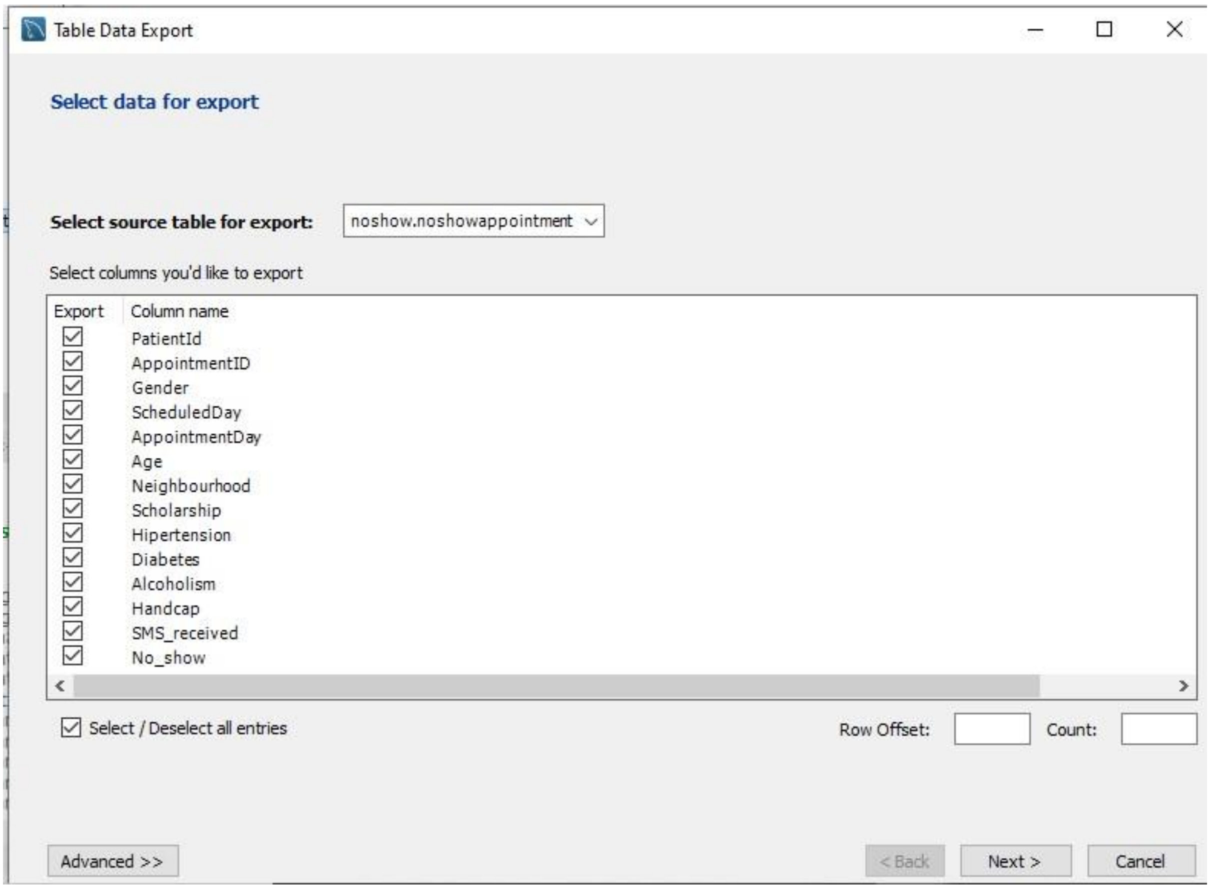
```
SMS_received, COUNT(*) AS NoShowCount
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Table: noshowappointments

Columns:
PatientId
AppointmentID
Gender
ScheduledDay
AppointmentDay
Age
Neighbourhood varchar
Scholarship varchar
Hipertension varchar
Diabetes varchar
Alcoholism varchar

```
 8 22:24:00 SELECT MIN(Age) from noshowappointments                                   1 row(s) returned    0.000 sec / 0.000 sec
 9 22:24:08 SELECT MIN(Age) from noshowappointments                                   1 row(s) returned    0.000 sec / 0.000 sec
10 22:24:14 SELECT MAX(Age) from noshowappointments                                   1 row(s) returned    0.000 sec / 0.000 sec
11 22:29:55 SELECT DISTINCT Neighbourhood FROM noshowappointments                     13 row(s) returned   0.000 sec / 0.000 sec
12 22:32:46 SELECT Gender, COUNT(*) AS NoShowCount FROM NoShowAppointments WHERE No_s... 2 row(s) returned    0.000 sec / 0.000 sec
```



Select data for export

Select source table for export: noshow.noshowappointment

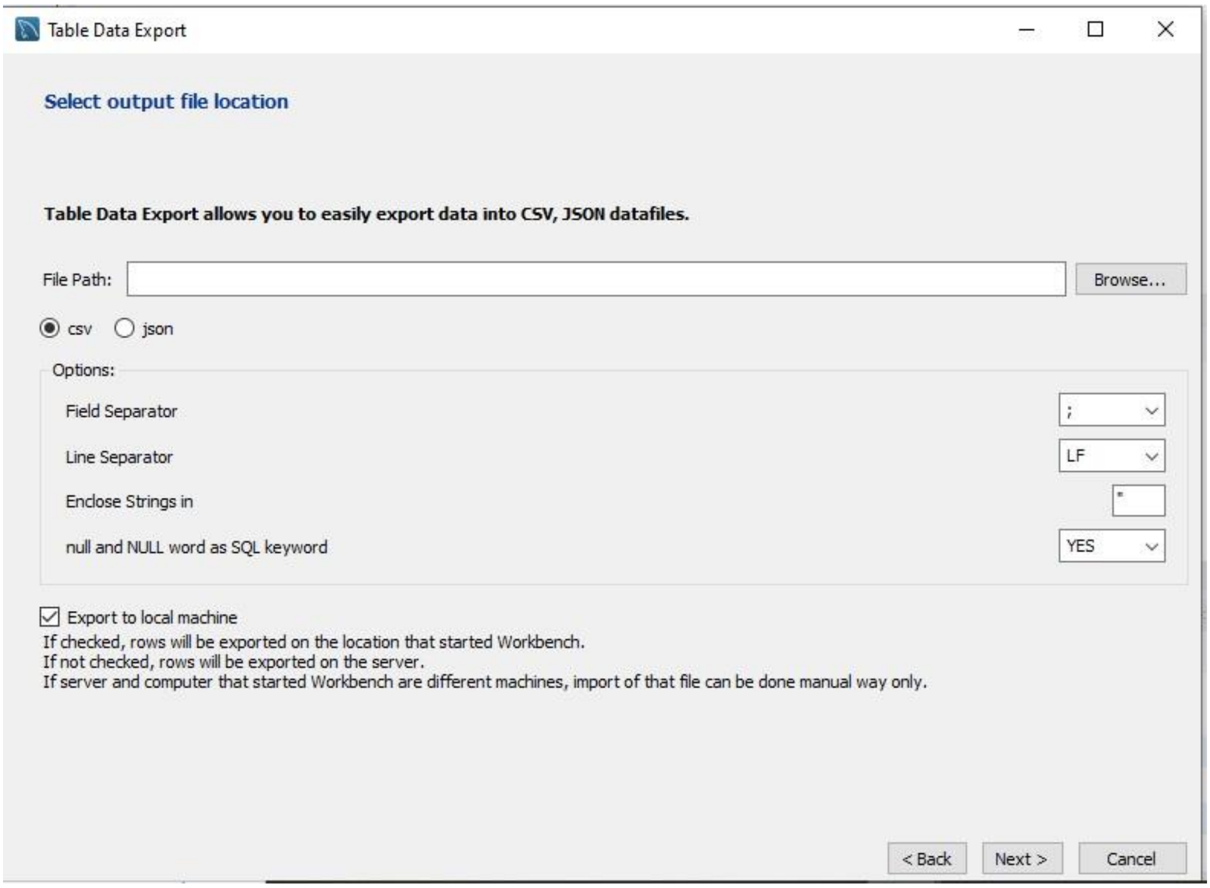
Select columns you'd like to export

Export	Column name
☑	PatientId
☑	AppointmentID
☑	Gender
☑	ScheduledDay
☑	AppointmentDay
☑	Age
☑	Neighbourhood
☑	Scholarship
☑	Hipertension
☑	Diabetes
☑	Alcoholism
☑	Handcap
☑	SMS_received
☑	No_show

☑ Select / Deselect all entries

Row Offset: Count:

Click **Next**



Select output file location

Table Data Export allows you to easily export data into CSV, JSON datafiles.

File Path:

● csv ○ json

Options:

Field Separator ;

Line Separator LF

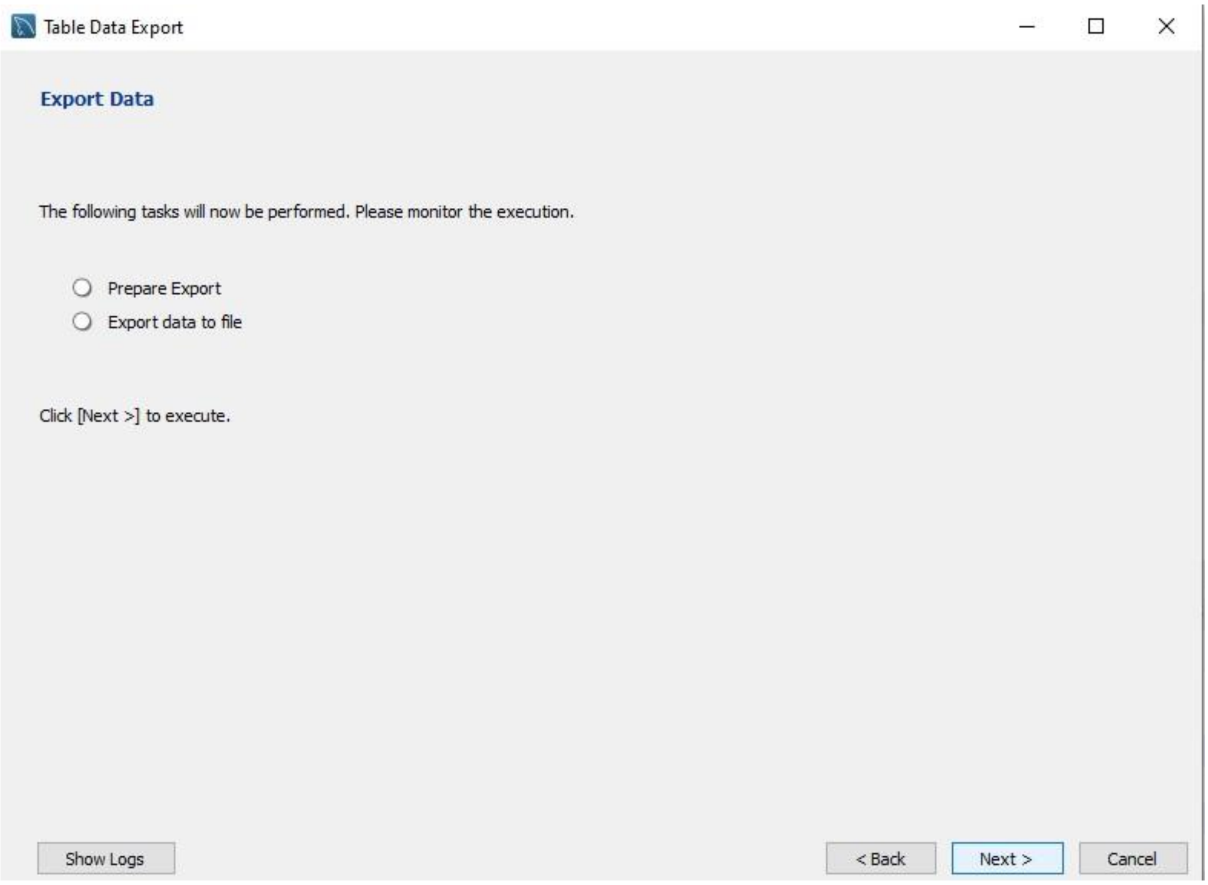
Enclose Strings in "

null and NULL word as SQL keyword YES

☑ Export to local machine

If checked, rows will be exported on the location that started Workbench.
If not checked, rows will be exported on the server.
If server and computer that started Workbench are different machines, import of that file can be done manual way only.

Click **Next** after selecting the folder and entering the name of the table to save.



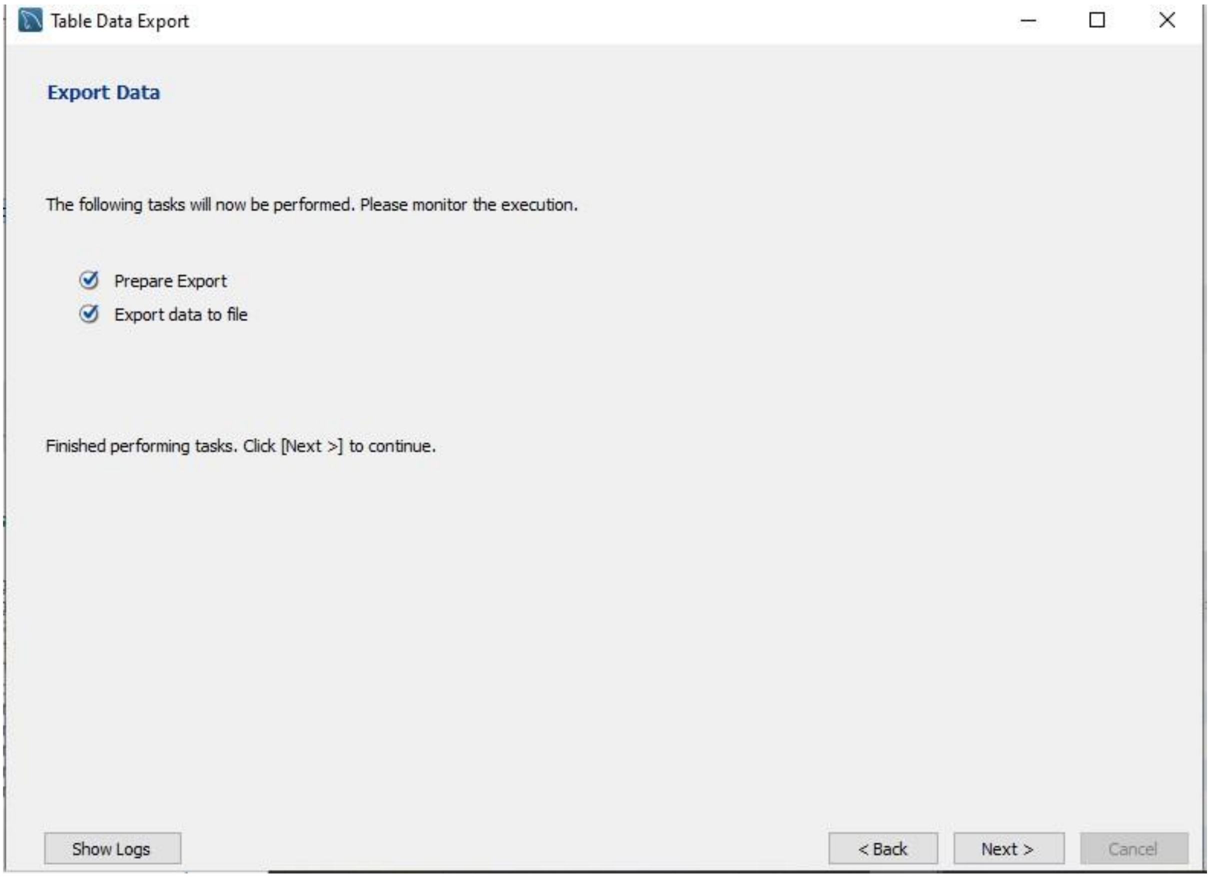
The following tasks will now be performed. Please monitor the execution.

○ Prepare Export

○ Export data to file

Click [Next >] to execute.

Click **Next**



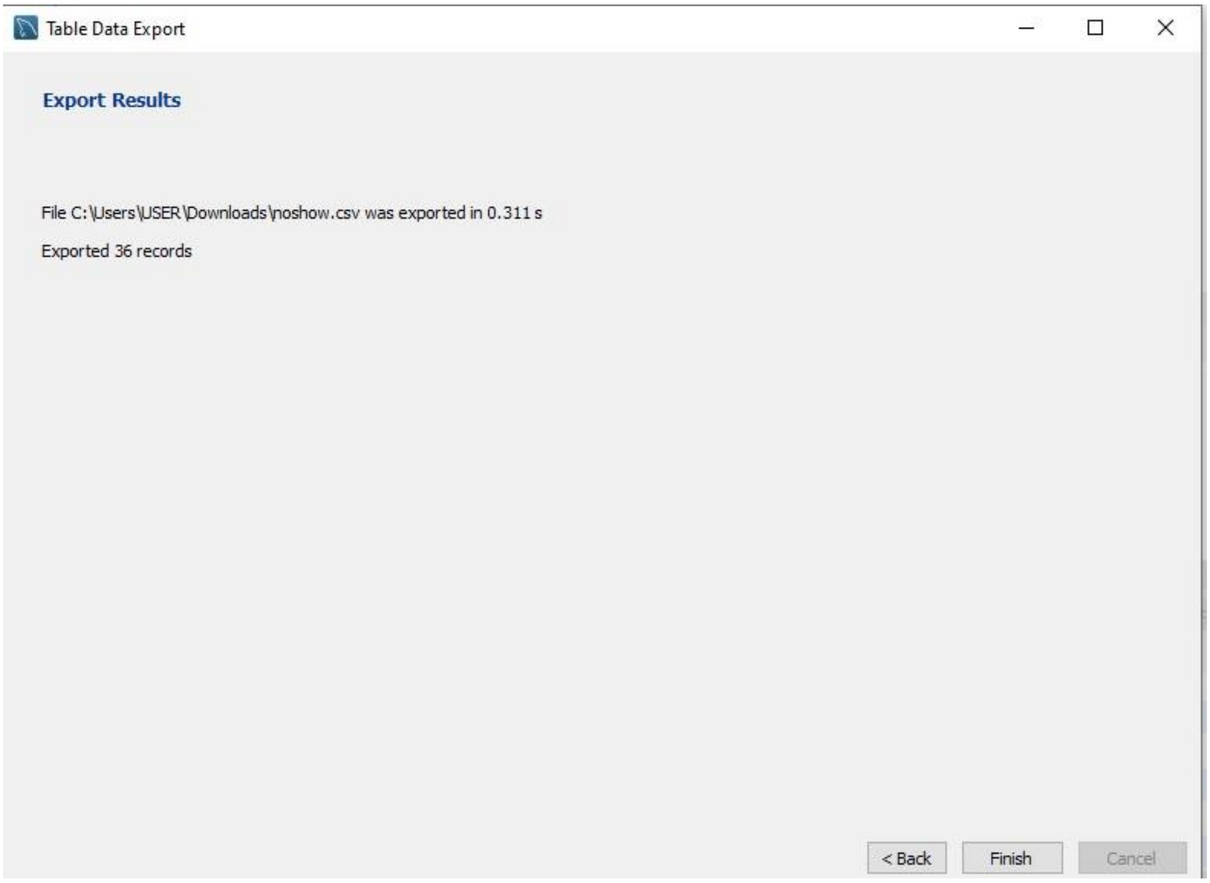
The following tasks will now be performed. Please monitor the execution.

✓ Prepare Export

✓ Export data to file

Finished performing tasks. Click [Next >] to continue.

Click **Finish** to finish exporting the table in csv format.



File C:\Users\USER\Downloads\noshow.csv was exported in 0.311 s

Exported 36 records

Open excel and click the **Data** tab and select **From Text/CSV**.

Select the CSV file exported from MySQL and click **Import.**

Then click **Load** then the data will be imported in excel.

Click **Save.**

- **Testing:** Test the integration and functionality of your Excel dashboard.

Done

Part 6: Presentation

- **Pitch Deck:** Develop a 10-slide PowerPoint presentation as taught in the entrepreneurship module

https://gamma.app/docs/Improving-Healthcare-Appointment-Attendance-A-Data-Driven-Approac-2p25qmm163mdof0?mode=doc