

This project which contains 2 files named Covid Deaths and Covid Vaccinations, which are both loaded on SSMS . Data types of relevant field were changed to integer to avoid later conversion during code execution.

The aim of this Project is to analyse the dataset related to deaths and vaccinations due to Covid-19.

Initially, Total deaths were analysed grouped by Location.

SQLQuery1.sql - LA...T1FD7P\imran (62))\*

Select Location, MAX(Total\_deaths) as TotalDeathCount  
From PortfolioProject..CovidDeaths  
--Where location like '%states%'  
Where continent is not null  
Group by location  
order by TotalDeathCount desc

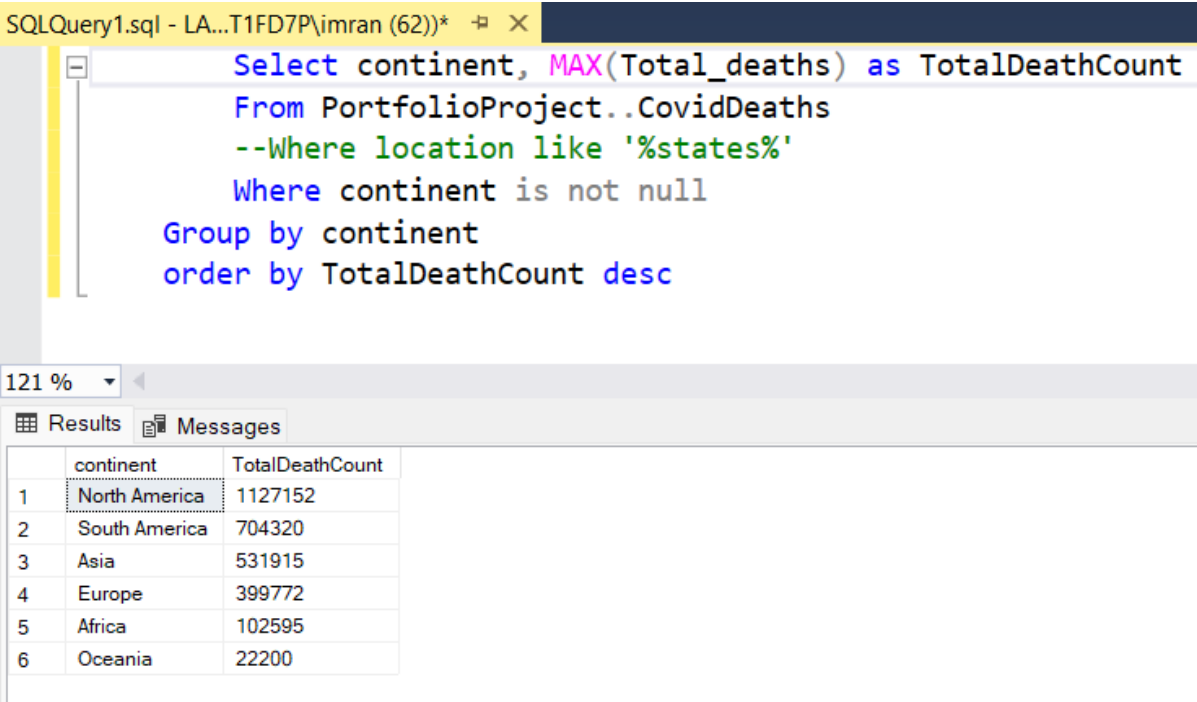
121 %

Results Messages

	Location	TotalDeathCount
1	United States	1127152
2	Brazil	704320
3	India	531915
4	Russia	399772
5	Mexico	334336
6	United Kingdom	228144
7	Peru	221203
8	Italy	190942
9	Germany	174979
10	France	167985
11	Indonesia	161879
12	Iran	146303
13	Colombia	142892
14	Argentina	130472
15	Spain	121852
16	China	121536
17	Poland	119632
18	Ukraine	109904
19	South Africa	102595
20	Turkey	101419
21	Japan	74694
22	Romania	68246
23	Philippines	66510
24	Chile	64497
25	Canada	53063
26	Hungary	48798

✓ Query executed successfully.

Then total deaths were analysed by continents using the syntax below.



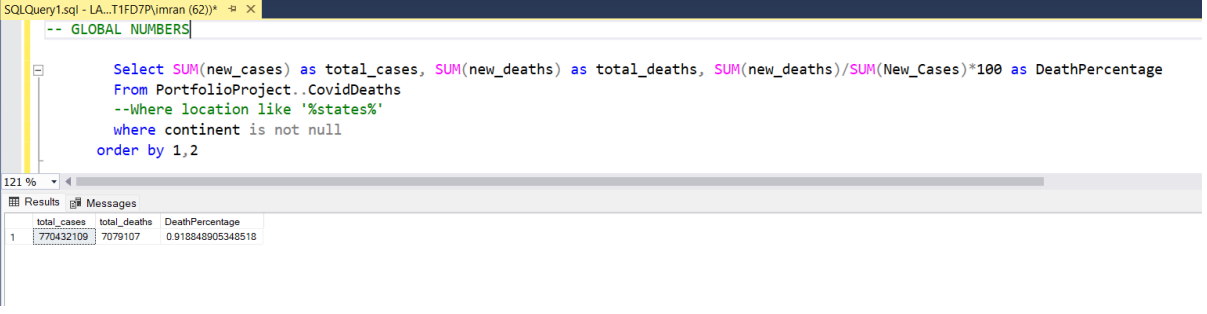
The screenshot shows a SQL query in a text editor window titled 'SQLQuery1.sql'. The query is as follows:

```
Select continent, MAX(Total_deaths) as TotalDeathCount
From PortfolioProject..CovidDeaths
--Where location like '%states%'
Where continent is not null
Group by continent
order by TotalDeathCount desc
```

Below the query editor, the 'Results' tab is active, displaying a table with two columns: 'continent' and 'TotalDeathCount'. The results are ordered by 'TotalDeathCount' in descending order.

	continent	TotalDeathCount
1	North America	1127152
2	South America	704320
3	Asia	531915
4	Europe	399772
5	Africa	102595
6	Oceania	22200

Next the global deaths percentage was queried



The screenshot shows a SQL query in a text editor window titled 'SQLQuery1.sql'. The query is as follows:

```
-- GLOBAL NUMBERS
Select SUM(new_cases) as total_cases, SUM(new_deaths) as total_deaths, SUM(new_deaths)/SUM(New_Cases)*100 as DeathPercentage
From PortfolioProject..CovidDeaths
--Where location like '%states%'
where continent is not null
order by 1,2
```

Below the query editor, the 'Results' tab is active, displaying a table with three columns: 'total\_cases', 'total\_deaths', and 'DeathPercentage'. The results are ordered by the first two columns.

	total_cases	total_deaths	DeathPercentage
1	770432109	7079107	0.918848905348518

Using the Syntax below, Percentage of population was analysed who received atleast one Covid Vaccination and this was accomplished by joining two tables covid deaths and covid vaccinations

SQLQuery1.sql - LA...T1FD7P(imran (62)) \* X

```
-- Shows Percentage of Population that has recieved at least one Covid Vaccine
```

```
Select dea.continent, dea.location, dea.date, dea.population, vac.new_vaccinations
, SUM(vac.new_vaccinations) OVER (Partition by dea.Location Order by dea.location, dea.Date) as RollingPeopleVaccinated
--, (RollingPeopleVaccinated/population)*100
From PortfolioProject..CovidDeaths dea
Join PortfolioProject..CovidVaccinations vac
On dea.location = vac.location
and dea.date = vac.date
where dea.continent is not null
order by 2,3
```

121 %

Results Messages

	continent	location	date	population	new_vaccinations	RollingPeopleVaccinated
1	Asia	Afghanistan	2020-01-03 00:00:00.000	41128772	NULL	NULL
2	Asia	Afghanistan	2020-01-04 00:00:00.000	41128772	NULL	NULL
3	Asia	Afghanistan	2020-01-05 00:00:00.000	41128772	NULL	NULL
4	Asia	Afghanistan	2020-01-06 00:00:00.000	41128772	NULL	NULL
5	Asia	Afghanistan	2020-01-07 00:00:00.000	41128772	NULL	NULL
6	Asia	Afghanistan	2020-01-08 00:00:00.000	41128772	NULL	NULL
7	Asia	Afghanistan	2020-01-09 00:00:00.000	41128772	NULL	NULL
8	Asia	Afghanistan	2020-01-10 00:00:00.000	41128772	NULL	NULL
9	Asia	Afghanistan	2020-01-11 00:00:00.000	41128772	NULL	NULL
10	Asia	Afghanistan	2020-01-12 00:00:00.000	41128772	NULL	NULL
11	Asia	Afghanistan	2020-01-13 00:00:00.000	41128772	NULL	NULL
12	Asia	Afghanistan	2020-01-14 00:00:00.000	41128772	NULL	NULL
13	Asia	Afghanistan	2020-01-15 00:00:00.000	41128772	NULL	NULL
14	Asia	Afghanistan	2020-01-16 00:00:00.000	41128772	NULL	NULL
15	Asia	Afghanistan	2020-01-17 00:00:00.000	41128772	NULL	NULL
16	Asia	Afghanistan	2020-01-18 00:00:00.000	41128772	NULL	NULL
17	Asia	Afghanistan	2020-01-19 00:00:00.000	41128772	NULL	NULL
18	Asia	Afghanistan	2020-01-20 00:00:00.000	41128772	NULL	NULL
19	Asia	Afghanistan	2020-01-21 00:00:00.000	41128772	NULL	NULL
20	Asia	Afghanistan	2020-01-22 00:00:00.000	41128772	NULL	NULL
21	Asia	Afghanistan	2020-01-23 00:00:00.000	41128772	NULL	NULL
22	Asia	Afghanistan	2020-01-24 00:00:00.000	41128772	NULL	NULL

Query executed successfully. LAPTOP-HIT1FD7P\SQLEXPRESS ... LAPTOP-HIT1FD7P\imran ... master 00:00:12 311,686 rows

Furthermore, analysing the mortality rate using new deaths vs population was carried out using CTE

SQLQuery1.sql - LAP...n (62)) Executing... \* X

```
--Using CTE to perform Calculation to perform MortalityRate (New deaths vs Population)
```

```
WITH table0 as
(Select dea.continent, dea.location, dea.population, vac.new_vaccinations, dea.new_cases, dea.new_deaths, dea.total_deaths,
(dea.new_deaths/dea.population)*100 as MortalityRate
from PortfolioProject..CovidDeaths dea
Join PortfolioProject..CovidVaccinations vac
On dea.location = vac.location
)
Select * from table0
where MortalityRate > 0
order by location
```

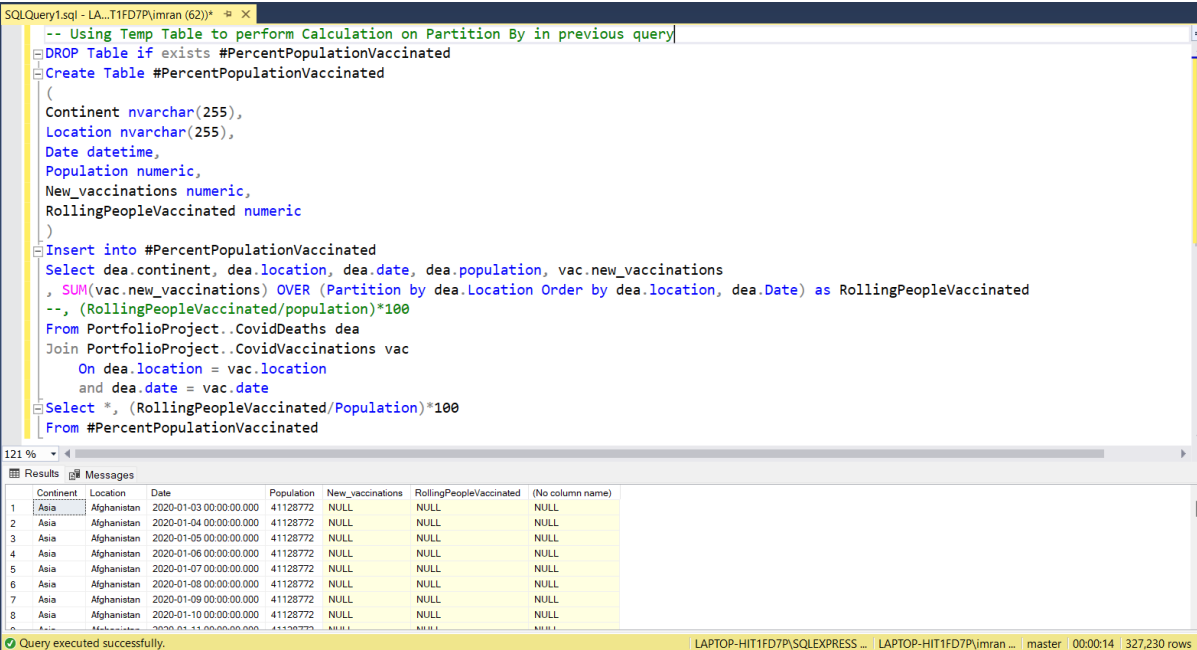
121 %

Results Messages

	continent	location	population	new_vaccinations	new_cases	new_deaths	total_deaths	MortalityRate
1	Asia	Afghanistan	41128772	NULL	18	1	7847	2.43138793446106E-06
2	Asia	Afghanistan	41128772	NULL	43	2	7849	4.86277586892213E-06
3	Asia	Afghanistan	41128772	NULL	52	1	7850	2.43138793446106E-06
4	Asia	Afghanistan	41128772	NULL	13	1	7851	2.43138793446106E-06
5	Asia	Afghanistan	41128772	NULL	30	1	7852	2.43138793446106E-06
6	Asia	Afghanistan	41128772	NULL	31	2	7854	4.86277586892213E-06
7	Asia	Afghanistan	41128772	NULL	25	1	7855	2.43138793446106E-06
8	Asia	Afghanistan	41128772	NULL	24	2	7857	4.86277586892213E-06
9	Asia	Afghanistan	41128772	NULL	16	2	7859	4.86277586892213E-06
10	Asia	Afghanistan	41128772	NULL	50	1	7860	2.43138793446106E-06
11	Asia	Afghanistan	41128772	NULL	125	1	7861	2.43138793446106E-06
12	Asia	Afghanistan	41128772	NULL	35	1	7862	2.43138793446106E-06
13	Asia	Afghanistan	41128772	NULL	74	1	7863	2.43138793446106E-06
14	Asia	Afghanistan	41128772	NULL	6	2	7865	4.86277586892213E-06
15	Asia	Afghanistan	41128772	NULL	3	1	7866	2.43138793446106E-06
16	Asia	Afghanistan	41128772	NULL	27	2	7868	4.86277586892213E-06
17	Asia	Afghanistan	41128772	NULL	27	1	7869	2.43138793446106E-06

Executing query... LAPTOP-HIT1FD7P\SQLEXPRESS ... LAPTOP-HIT1FD7P\imran ... master 00:00:12 311,686 rows

Temp table was created



The screenshot shows a SQL Server Enterprise Manager interface. The top pane displays a SQL query window with the following code:

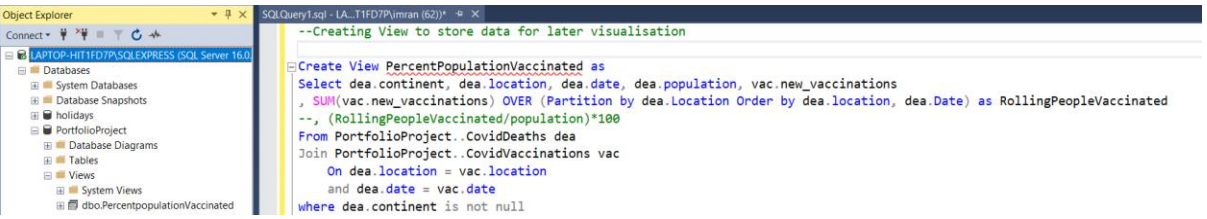
```
-- Using Temp Table to perform Calculation on Partition By in previous query
DROP Table if exists #PercentPopulationVaccinated
Create Table #PercentPopulationVaccinated
(
    Continent nvarchar(255),
    Location nvarchar(255),
    Date datetime,
    Population numeric,
    New_vaccinations numeric,
    RollingPeopleVaccinated numeric
)
Insert into #PercentPopulationVaccinated
Select dea.continent, dea.location, dea.date, dea.population, vac.new_vaccinations
, SUM(vac.new_vaccinations) OVER (Partition by dea.Location Order by dea.location, dea.Date) as RollingPeopleVaccinated
--, (RollingPeopleVaccinated/population)*100
From PortfolioProject..CovidDeaths dea
Join PortfolioProject..CovidVaccinations vac
    On dea.location = vac.location
    and dea.date = vac.date
Select *, (RollingPeopleVaccinated/Population)*100
From #PercentPopulationVaccinated
```

The bottom pane shows the results of the query, displaying a table with 8 rows and 7 columns. The columns are: Continent, Location, Date, Population, New\_vaccinations, RollingPeopleVaccinated, and (No column name). The data is as follows:

	Continent	Location	Date	Population	New_vaccinations	RollingPeopleVaccinated	(No column name)
1	Asia	Afghanistan	2020-01-03 00:00:00.000	41128772	NULL	NULL	NULL
2	Asia	Afghanistan	2020-01-04 00:00:00.000	41128772	NULL	NULL	NULL
3	Asia	Afghanistan	2020-01-05 00:00:00.000	41128772	NULL	NULL	NULL
4	Asia	Afghanistan	2020-01-06 00:00:00.000	41128772	NULL	NULL	NULL
5	Asia	Afghanistan	2020-01-07 00:00:00.000	41128772	NULL	NULL	NULL
6	Asia	Afghanistan	2020-01-08 00:00:00.000	41128772	NULL	NULL	NULL
7	Asia	Afghanistan	2020-01-09 00:00:00.000	41128772	NULL	NULL	NULL
8	Asia	Afghanistan	2020-01-10 00:00:00.000	41128772	NULL	NULL	NULL

The status bar at the bottom indicates "Query executed successfully." and "LAPTOP-HIT1FD7P\SQLEXPRESS ... LAPTOP-HIT1FD7P\imran ... master | 00:00:14 | 327,230 rows".

Finally to store the data for later visualisation, Create view was created



The screenshot shows a SQL Server Enterprise Manager interface. The top pane displays a SQL query window with the following code:

```
--Creating View to store data for later visualisation
Create View PercentPopulationVaccinated as
Select dea.continent, dea.location, dea.date, dea.population, vac.new_vaccinations
, SUM(vac.new_vaccinations) OVER (Partition by dea.Location Order by dea.location, dea.Date) as RollingPeopleVaccinated
--, (RollingPeopleVaccinated/population)*100
From PortfolioProject..CovidDeaths dea
Join PortfolioProject..CovidVaccinations vac
    On dea.location = vac.location
    and dea.date = vac.date
where dea.continent is not null
```

The bottom pane shows the Object Explorer, which displays the database structure. The database is named "LAPTOP-HIT1FD7P\SQLEXPRESS (SQL Server 16.0)". The database contains several system databases, snapshots, and a "PortfolioProject" database. The "PortfolioProject" database contains a "dbo" schema, which contains a view named "PercentPopulationVaccinated".