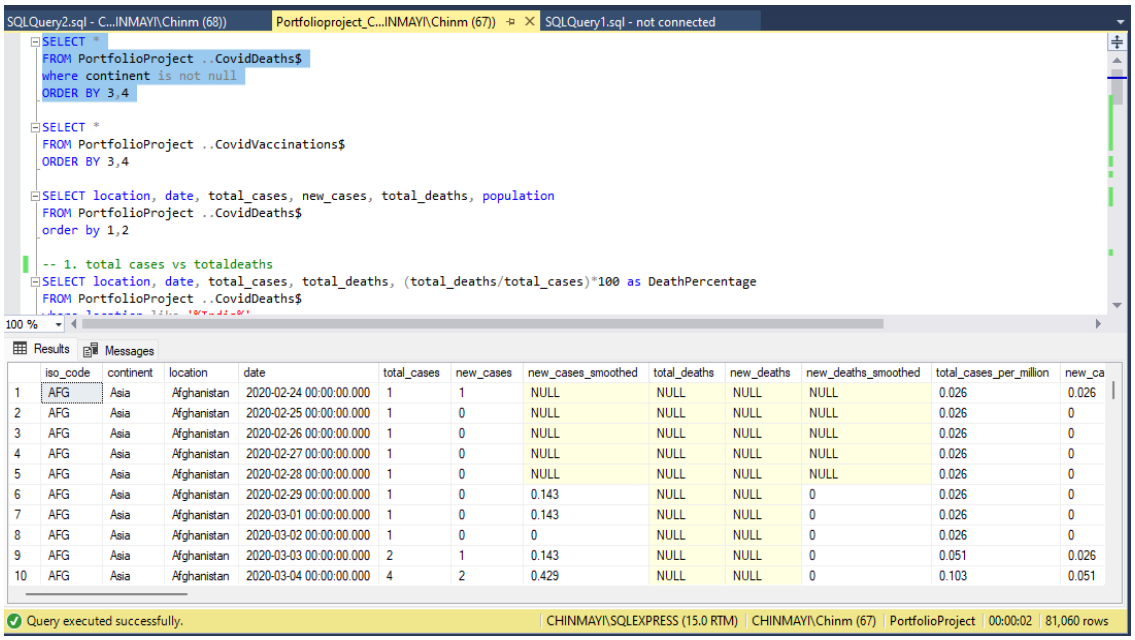
SQL PORTFOLIO PROJECT – COVID 19 ANALYSIS REPORT

(S.CHINMAYI)

This is a Covid-19 Analysis Project report consisting of the Queries performed and the Output obtained. The RDBMS Tool used for this Project is SSMS.

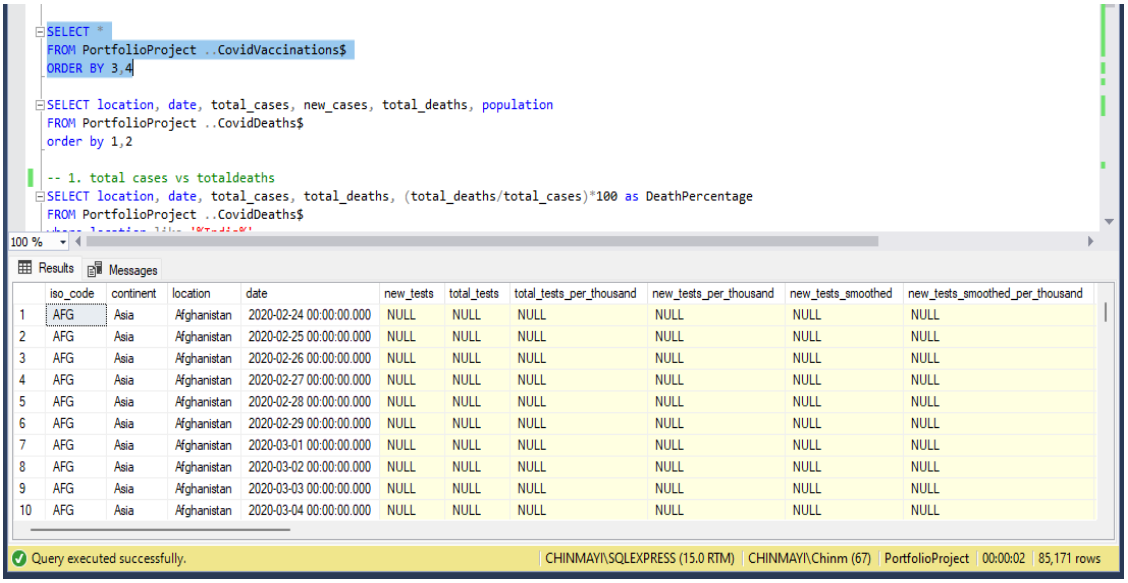
The two tables used for this project are CovidDeaths and CovidVaccinations



```
SELECT *
FROM PortfolioProject ..CovidDeaths$
where continent is not null
ORDER BY 3,4

SELECT *
FROM PortfolioProject ..CovidVaccinations$
ORDER BY 3,4

SELECT location, date, total_cases, new_cases, total_deaths, population
FROM PortfolioProject ..CovidDeaths$
order by 1,2

-- 1. total cases vs totaldeaths
SELECT location, date, total_cases, total_deaths, (total_deaths/total_cases)*100 as DeathPercentage
FROM PortfolioProject ..CovidDeaths$
```

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_deaths	new_deaths	new_deaths_smoothed	total_cases_per_million	new_ca
1	AFG	Asia	Afghanistan	2020-02-24 00:00:00.000	1	1	NULL	NULL	NULL	NULL	0.026	0.026
2	AFG	Asia	Afghanistan	2020-02-25 00:00:00.000	1	0	NULL	NULL	NULL	NULL	0.026	0
3	AFG	Asia	Afghanistan	2020-02-26 00:00:00.000	1	0	NULL	NULL	NULL	NULL	0.026	0
4	AFG	Asia	Afghanistan	2020-02-27 00:00:00.000	1	0	NULL	NULL	NULL	NULL	0.026	0
5	AFG	Asia	Afghanistan	2020-02-28 00:00:00.000	1	0	NULL	NULL	NULL	NULL	0.026	0
6	AFG	Asia	Afghanistan	2020-02-29 00:00:00.000	1	0	0.143	NULL	NULL	0	0.026	0
7	AFG	Asia	Afghanistan	2020-03-01 00:00:00.000	1	0	0.143	NULL	NULL	0	0.026	0
8	AFG	Asia	Afghanistan	2020-03-02 00:00:00.000	1	0	0	NULL	NULL	0	0.026	0
9	AFG	Asia	Afghanistan	2020-03-03 00:00:00.000	2	1	0.143	NULL	NULL	0	0.051	0.026
10	AFG	Asia	Afghanistan	2020-03-04 00:00:00.000	4	2	0.429	NULL	NULL	0	0.103	0.051

Query executed successfully. CHINMAYI\SQLEXPRESS (15.0 RTM) CHINMAYI\Chinm (67) PortfolioProject 00:00:02 81,060 rows



```
SELECT *
FROM PortfolioProject ..CovidVaccinations$
ORDER BY 3,4

SELECT location, date, total_cases, new_cases, total_deaths, population
FROM PortfolioProject ..CovidDeaths$
order by 1,2

-- 1. total cases vs totaldeaths
SELECT location, date, total_cases, total_deaths, (total_deaths/total_cases)*100 as DeathPercentage
FROM PortfolioProject ..CovidDeaths$
```

	iso_code	continent	location	date	new_tests	total_tests	total_tests_per_thousand	new_tests_per_thousand	new_tests_smoothed	new_tests_smoothed_per_thousand
1	AFG	Asia	Afghanistan	2020-02-24 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
2	AFG	Asia	Afghanistan	2020-02-25 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
3	AFG	Asia	Afghanistan	2020-02-26 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
4	AFG	Asia	Afghanistan	2020-02-27 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
5	AFG	Asia	Afghanistan	2020-02-28 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
6	AFG	Asia	Afghanistan	2020-02-29 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
7	AFG	Asia	Afghanistan	2020-03-01 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
8	AFG	Asia	Afghanistan	2020-03-02 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
9	AFG	Asia	Afghanistan	2020-03-03 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL
10	AFG	Asia	Afghanistan	2020-03-04 00:00:00.000	NULL	NULL	NULL	NULL	NULL	NULL

Query executed successfully. CHINMAYI\SQLEXPRESS (15.0 RTM) CHINMAYI\Chinm (67) PortfolioProject 00:00:02 85,171 rows

1.Total Cases vs Total Deaths

```
order by 1,2

-- 1. total cases vs totaldeaths
SELECT location, date, total_cases, total_deaths, (total_deaths/total_cases)*100 as DeathPercentage
FROM PortfolioProject ..CovidDeaths$
where location like '%India%'
order by 1,2

-- 2.Totalcases vs population
SELECT location, date, total_cases, population, (total_cases/population)*100 as DeathPercentage
FROM PortfolioProject ..CovidDeaths$
where location like '%India%'
```

100 %

Results | Messages

	location	date	total_cases	total_deaths	DeathPercentage
1	India	2020-01-30 00:00:00.000	1	NULL	NULL
2	India	2020-01-31 00:00:00.000	1	NULL	NULL
3	India	2020-02-01 00:00:00.000	1	NULL	NULL
4	India	2020-02-02 00:00:00.000	2	NULL	NULL
5	India	2020-02-03 00:00:00.000	3	NULL	NULL
6	India	2020-02-04 00:00:00.000	3	NULL	NULL
7	India	2020-02-05 00:00:00.000	3	NULL	NULL
8	India	2020-02-06 00:00:00.000	3	NULL	NULL
9	India	2020-02-07 00:00:00.000	3	NULL	NULL
10	India	2020-02-08 00:00:00.000	3	NULL	NULL
11	India	2020-02-09 00:00:00.000	3	NULL	NULL

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:00 | 457 rows

We get an output of about 210 rows after executing the query.

2.Total Cases vs Total Population

```
-- 2.Totalcases vs population
SELECT location, date, total_cases, population, (total_cases/population)*100 as DeathPercentage
FROM PortfolioProject ..CovidDeaths$
where location like '%India%'
order by 1,2

-- 3.countries which have highest infection rate compared to population
SELECT location, Population, MAX(total_cases) as HighestInfectionCount, MAX((total_cases/population))*100 as PercentPopulationInfected
FROM PortfolioProject ..CovidDeaths$
  --where location like '%India%'
Group by Location, Population
```

100 %

Results | Messages

	location	date	total_cases	population	DeathPercentage
1	India	2020-01-30 00:00:00.000	1	1380004385	7.2463537860425E-08
2	India	2020-01-31 00:00:00.000	1	1380004385	7.2463537860425E-08
3	India	2020-02-01 00:00:00.000	1	1380004385	7.2463537860425E-08
4	India	2020-02-02 00:00:00.000	2	1380004385	1.4492707572085E-07
5	India	2020-02-03 00:00:00.000	3	1380004385	2.17390613581275E-07
6	India	2020-02-04 00:00:00.000	3	1380004385	2.17390613581275E-07
7	India	2020-02-05 00:00:00.000	3	1380004385	2.17390613581275E-07
8	India	2020-02-06 00:00:00.000	3	1380004385	2.17390613581275E-07
9	India	2020-02-07 00:00:00.000	3	1380004385	2.17390613581275E-07
10	India	2020-02-08 00:00:00.000	3	1380004385	2.17390613581275E-07
11	India	2020-02-09 00:00:00.000	3	1380004385	2.17390613581275E-07

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:00 | 457 rows

3.Countries having Highest Infection Rates compared to population

```sql
-- 3.countries which have highest infection rate compared to population
SELECT location, Population, MAX(total_cases) as HighestInfectionCount, MAX((total_cases/population))*100 as PercentPopulationInfected
FROM PortfolioProject ..CovidDeaths$
  --where location like '%India%'
Group by Location, Population
order by PercentPopulationInfected desc

-- 4.countries which have highest death count per population
SELECT location, MAX(cast(total_deaths as int)) as TotalDeathCount
FROM PortfolioProject ..CovidDeaths$
  --where location like '%India%
where continent is not null
Group by Location, Population
```

Results | Messages

	location	Population	HighestInfectionCount	PercentPopulationInfected
1	Andorra	77265	13232	17.1254772536077
2	Montenegro	628062	97389	15.5062716738156
3	Czechia	10708982	1630758	15.2279460363273
4	San Marino	33938	5066	14.9272202251164
5	Slovenia	2078932	240292	11.5584348117206
6	Luxembourg	625976	67205	10.7360346083556
7	Bahrain	1701583	176934	10.3981997939566
8	Serbia	6804596	689557	10.1336949320724
9	United States	331002647	32346971	9.77242064169958
10	Israel	8655541	838481	9.68721654718059
11	Sweden	10099270	973604	9.64034034143062

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:00 | 219 rows

Andorra has Highest Infection Rate of about 17.12%.

4.Countries having Highest Death Count per Population

```sql
-- 4.countries which have highest death count per population
SELECT location, MAX(cast(total_deaths as int)) as TotalDeathCount
FROM PortfolioProject ..CovidDeaths$
  --where location like '%India%
where continent is not null
Group by Location, Population
order by TotalDeathCount desc

-- 5.BREAK DOWN THE DATA BY CONTINENTS
SELECT continent, MAX(cast(total_deaths as int)) as TotalDeathCount
FROM PortfolioProject ..CovidDeaths$
where continent is not null
Group by continent
order by TotalDeathCount desc
```

Results | Messages

	location	TotalDeathCount
1	United States	576232
2	Brazil	403781
3	Mexico	216907
4	India	211853
5	United Kingdom	127775
6	Italy	120807
7	Russia	108290
8	France	104675
9	Germany	83097
10	Spain	78216
11	Colombia	73720

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:00 | 210 rows

United States has the Highest Death Count

5. Analyzing the data Continent-wise

```
-- 5.BREAK DOWN THE DATA BY CONTINENTS
SELECT continent, MAX(cast(total_deaths as int)) as TotalDeathCount
FROM PortfolioProject ..CovidDeaths$
where continent is not null
Group by continent
order by TotalDeathCount desc

-- 6.Global Numbers
SELECT date, SUM(new_cases) as total_cases, SUM(cast(new_deaths as int)) as total_deaths, SUM(cast(new_deaths as int))/SUM(New_Cases)*100 as DeathPe
FROM PortfolioProject..CovidDeaths$
where continent is not null
GROUP BY date
order by 1,2
```

00 %

Results | Messages

	continent	TotalDeathCount
1	North America	576232
2	South America	403781
3	Asia	211853
4	Europe	127775
5	Africa	54350
6	Oceania	910

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:00 | 6 rows

```
-- 6.Global Numbers
SELECT date, SUM(new_cases) as total_cases, SUM(cast(new_deaths as int)) as total_deaths, SUM(cast(new_deaths as int))/SUM(New_Cases)*100 as DeathPe
FROM PortfolioProject..CovidDeaths$
where continent is not null
GROUP BY date
order by 1,2

-- 7.Death percent across the world
SELECT SUM(new_cases) as total_cases, SUM(cast(new_deaths as int)) as total_deaths, SUM(cast(new_deaths as int))/SUM(New_Cases)*100 as DeathPercenta
FROM PortfolioProject..CovidDeaths$
where continent is not null
order by 1,2
```

100 %

Results | Messages

	date	total_cases	total_deaths	DeathPercentage
1	2020-01-01 00:00:00.000	NULL	NULL	NULL
2	2020-01-02 00:00:00.000	NULL	NULL	NULL
3	2020-01-03 00:00:00.000	NULL	NULL	NULL
4	2020-01-04 00:00:00.000	NULL	NULL	NULL
5	2020-01-05 00:00:00.000	NULL	NULL	NULL
6	2020-01-06 00:00:00.000	NULL	NULL	NULL
7	2020-01-07 00:00:00.000	NULL	NULL	NULL
8	2020-01-08 00:00:00.000	NULL	NULL	NULL
9	2020-01-09 00:00:00.000	NULL	NULL	NULL
10	2020-01-10 00:00:00.000	NULL	NULL	NULL
11	2020-01-11 00:00:00.000	NULL	NULL	NULL

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:00 | 486 rows

5.Death Percent Across the World

```sql
-- 7.Death percent across the world
SELECT SUM(new_cases) as total_cases, SUM(cast(new_deaths as int))
FROM PortfolioProject..CovidDeaths$
where continent is not null
order by 1,2

-- 8.Looking at Total Population vs Vaccinations
SELECT dea.continent, dea.location, dea.date,dea.population, vac.n
      (Partition by dea.location ORDER BY dea.location, dea.date)
      -- , (RollingPeopleVaccinated/population)*100
FROM PortfolioProject..CovidDeaths$ as dea
Join PortfolioProject..CovidVaccinations$ as vac
      on dea.location = vac.location
      and dea.date = vac.date
where dea.continent is not null
```

Results | Messages

total_cases	total_deaths	DeathPercentage
150574977	3180206	2.11204149810363

6.Total Population vs Vaccinations

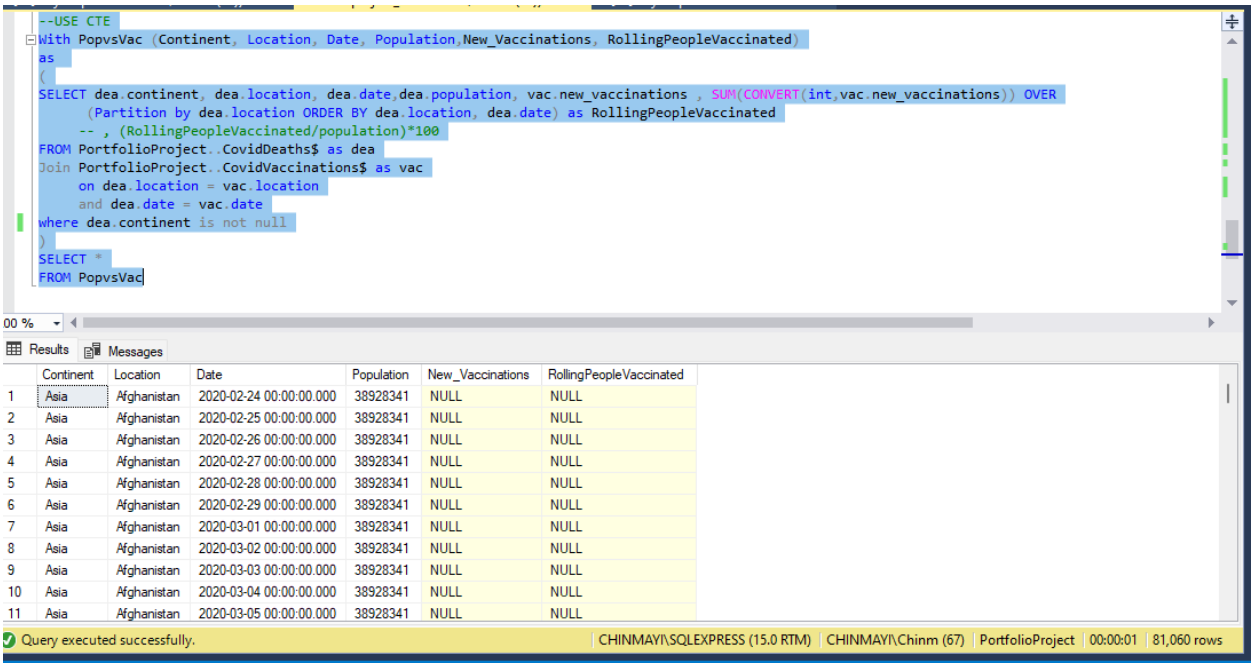
```sql
-- 8.Looking at Total Population vs Vaccinations
SELECT dea.continent, dea.location, dea.date,dea.population, vac.new_vaccinations , SUM(CONVERT(int,vac.new_vaccinations)) OVER
      (Partition by dea.location ORDER BY dea.location, dea.date) as RollingPeopleVaccinated
      -- , (RollingPeopleVaccinated/population)*100
FROM PortfolioProject..CovidDeaths$ as dea
Join PortfolioProject..CovidVaccinations$ as vac
      on dea.location = vac.location
      and dea.date = vac.date
where dea.continent is not null
order by  2, 3

--USE CTE
With PopvsVac (Continent, Location, Date, Population,New_Vaccinations, RollingPeopleVaccinated)
as
```

100 %

Results | Messages

	continent	location	date	population	new_vaccinations	RollingPeopleVaccinated
1	Asia	Afghanistan	2020-02-24 00:00:00.000	38928341	NULL	NULL
2	Asia	Afghanistan	2020-02-25 00:00:00.000	38928341	NULL	NULL
3	Asia	Afghanistan	2020-02-26 00:00:00.000	38928341	NULL	NULL
4	Asia	Afghanistan	2020-02-27 00:00:00.000	38928341	NULL	NULL
5	Asia	Afghanistan	2020-02-28 00:00:00.000	38928341	NULL	NULL
6	Asia	Afghanistan	2020-02-29 00:00:00.000	38928341	NULL	NULL
7	Asia	Afghanistan	2020-03-01 00:00:00.000	38928341	NULL	NULL
8	Asia	Afghanistan	2020-03-02 00:00:00.000	38928341	NULL	NULL
9	Asia	Afghanistan	2020-03-03 00:00:00.000	38928341	NULL	NULL
10	Asia	Afghanistan	2020-03-04 00:00:00.000	38928341	NULL	NULL
11	Asia	Afghanistan	2020-03-05 00:00:00.000	38928341	NULL	NULL

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:01 | 81,060 rows



```sql
--USE CTE
with PopvsVac (Continent, Location, Date, Population,New_Vaccinations, RollingPeopleVaccinated)
as
(
SELECT dea.continent, dea.location, dea.date,dea.population, vac.new_vaccinations , SUM(CONVERT(int,vac.new_vaccinations)) OVER
     (Partition by dea.location ORDER BY dea.location, dea.date) as RollingPeopleVaccinated
    -- , (RollingPeopleVaccinated/population)*100
FROM PortfolioProject..CovidDeaths$ as dea
Join PortfolioProject..CovidVaccinations$ as vac
    on dea.location = vac.location
    and dea.date = vac.date
where dea.continent is not null
)
SELECT *
FROM PopvsVac
```

00 %

Results | Messages

	Continent	Location	Date	Population	New_Vaccinations	RollingPeopleVaccinated
1	Asia	Afghanistan	2020-02-24 00:00:00.000	38928341	NULL	NULL
2	Asia	Afghanistan	2020-02-25 00:00:00.000	38928341	NULL	NULL
3	Asia	Afghanistan	2020-02-26 00:00:00.000	38928341	NULL	NULL
4	Asia	Afghanistan	2020-02-27 00:00:00.000	38928341	NULL	NULL
5	Asia	Afghanistan	2020-02-28 00:00:00.000	38928341	NULL	NULL
6	Asia	Afghanistan	2020-02-29 00:00:00.000	38928341	NULL	NULL
7	Asia	Afghanistan	2020-03-01 00:00:00.000	38928341	NULL	NULL
8	Asia	Afghanistan	2020-03-02 00:00:00.000	38928341	NULL	NULL
9	Asia	Afghanistan	2020-03-03 00:00:00.000	38928341	NULL	NULL
10	Asia	Afghanistan	2020-03-04 00:00:00.000	38928341	NULL	NULL
11	Asia	Afghanistan	2020-03-05 00:00:00.000	38928341	NULL	NULL

Query executed successfully. | CHINMAYI\SQLEXPRESS (15.0 RTM) | CHINMAYI\Chinm (67) | PortfolioProject | 00:00:01 | 81,060 rows