This analysis's goal is to gain insights with Exploratory Data Analys is applied to Covid-19 Data so that we can gain more detailed knowledge to understand Covid 19 better and how to prevent it in the future.This analysis is done in Google BigQuery with dataset imported from **www.ourworldindata.org.**

Total cases vs Total Death in each country. In this analysis,I will use Afghanistan, United States, China, Indonesia, and Singapore

```
SELECT
location,
date,
population,
total_cases,
total_deaths,
(total_deaths/total_cases) * 100 as Percentage_of_death
from
`covidanalysis-369714.COVID_DATASET.Covid_Death` where location in ('Afghanistan','United States','China','Indonesia','Singapore')
order by 2 desc limit 5
```

Total cases vs population in each country. To show the percentage of people affected by COVID-19. I use Afghanistan,United States, China, Indonesia, Singapore for the analysis

```
SELECT
location,
date,
population,
total_cases,
(total_cases/population) * 100 as Affected_Percentage
from
`covidanalysis-369714.COVID_DATASET.Covid_Death` where location in ('Afghanistan','United States','China','Indonesia','Singapore')
order by 2 desc limit 5
```

Countries with highest death rates

```
9SELECT
location,
max(total_deaths) as total_deaths,
population,
max((total_deaths/population) * 100) as Death_Rates_Percentage
from
```

`covidanalysis-369714.COVID_DATASET.Covid_Death` group by location,population order by Death_Rates_Percentage desc limit 10

Continents with highest death rate

```
select
continent,
max(total_deaths) as total_deaths
from `covidanalysis-369714.COVID_DATASET.Covid_Death` where continent is not null group by continent order by total_deaths desc
```

Global numbers

```
select  sum(new_cases),sum(new_deaths) as total_deaths,
sum(new_deaths)/sum(new_cases)*100 as DeathPercentage
from `covidanalysis-369714.COVID_DATASET.Covid_Death` where continent is not null
```

Counting vaccinations rolling per country. In this case I choose the vaccinations rollout of Indonesia.

```
select
dea.continent,
dea.location,
dea.date,
dea.population,
cast(vac.new_vaccinations as int) as new_vaccinations,
sum(cast(vac.new_vaccinations as int)) over (partition by dea.location order by dea.location,dea.date)
as Rolling_New_Vaccinations
from
`covidanalysis-369714.COVID_DATASET.Covid_Death` dea
inner join
`covidanalysis-369714.COVID_DATASET.Covid_vaccinations` vac
on dea.location = vac.location and dea.date= vac.date where dea.continent is not null
and dea.location = "Indonesia"
order by 2,3
```

Comparing vaccinations rollout to icu patients admissions across the country. The country in that I use in this analysis is Italy because of the severity of the covid cases in the country in the beginning of the pandemic

```
select
dea.continent,
dea.location,
dea.date,
dea.population,
cast(vac.new_vaccinations as int) as new_vaccinations,
sum(cast(vac.new_vaccinations as int)) over (partition by dea.location order by dea.location,dea.date) as
Rolling_New_Vaccinations,
cast(dea.icu_patients as int) as icu_patients_admissions,
sum(cast(dea.icu_patients as int)) over (partition by dea.location order by dea.location,dea.date) as
Rolling_New_ICU_Admissions
from
`covidanalysis-369714.COVID_DATASET.Covid_Death` dea
inner join
`covidanalysis-369714.COVID_DATASET.Covid_vaccinations` vac
on dea.location = vac.location and dea.date= vac.date where dea.continent is not null and
dea.location ="Italy"
order by 2,3
```

Analyzing the effects of diabetes prevalence number,handwashing facilities, and gdp per capita to covid infection rate

```
with temp_table as (SELECT
dea.location,
max(dea.total_cases) as total_cases,
dea.population,
max((dea.total_cases/dea.population) * 100) as Infection_Rate,
vac.diabetes_prevalence,
vac.handwashing_facilities,
vac.gdp_per_capita
from
`covidanalysis-369714.COVID_DATASET.Covid_Death` dea inner join
`covidanalysis-369714.COVID_DATASET.Covid_vaccinations` vac
on
dea.location = vac.location
where diabetes_prevalence is not null and vac.handwashing_facilities is not null and dea.location
!="World"
group by dea.location,dea.population
,vac.diabetes_prevalence,vac.handwashing_facilities,vac.gdp_per_capita
order by total_cases desc,Infection_Rate desc
)
```

Calculating correlation between infection rate and the 3 variables

```
select
```

```
corr(temp_table.Infection_Rate,diabetes_prevalence) as corr_diabetes,
corr(temp_table.Infection_Rate,handwashing_facilities) as corr_handwashing,
corr(temp_table.Infection_Rate,gdp_per_capita) as corr_gdp
from temp_table;
```