

```

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import matplotlib.pyplot as plt
import seaborn as sns
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
path="../../input/covid19-daily-data-updated/Covid-19_data_version_8.csv"
covid_data=pd.read_csv(path, parse_dates=True)
#vaccine_data.tail()
covid_data=covid_data.sort_values(by='date')

covid_data.people_vaccinated_per_hundred=covid_data.people_vaccinated_p
er_hundred/10
covid_data=covid_data.rename(columns={'people_vaccinated_per_hundred': '
people_vaccinated_per_thousand'})

covid_data.columns
high=covid_data.groupby('location').total_cases.agg([max]).sort_values(
by='max', ascending=False)
high.iloc[:18, :]
deaths=pd.DataFrame({'total_deaths':covid_data.groupby('continent').new
_deaths.sum()})
deaths
cases=pd.DataFrame({'total_cases':covid_data.groupby('continent').new_c
ases.sum()})
cases
cases_per_mil=covid_data.groupby('continent').total_cases_per_million.m
ean()
cases_per_mil
#covid_data.iloc[112463,0:33]

covid_data.female_smokers=covid_data.female_smokers.fillna(0)
covid_data.male_smokers=covid_data.male_smokers.fillna(0)
male=covid_data.male_smokers.notnull()
covid_data.female_smokers=covid_data.female_smokers+covid_data.male_smo
kers

covid_data=covid_data.rename(columns={'female_smokers': 'smokers'})
covid_data.drop('male_smokers', axis=1, inplace=True)
covid_data.people_vaccinated_per_thousand
covid_data_asia=covid_data.continent=='Asia'
plt.figure(figsize=(20,10))
#covid_data_cont.isnull=0
plt.title('Spread of covid-19 in different Continents')

sns.barplot(x=cases.index, y= cases.total_cases)
plt.legend()
covid_data_asia=covid_data.continent=='Asia'

```

```

plt.figure(figsize=(20,10))
# covid_data_cont.isnull=0
plt.title('Total deaths due to covid-19 in different Continents')

sns.barplot(x=deaths.index , y= deaths.total_deaths)
low=np.arange(0,35,0.1)
covid_data.smokers=covid_data.smokers.astype(int)
covid_data.smokers=covid_data['smokers'].replace(np.arange(0,25), "Low")
covid_data.smokers=covid_data['smokers'].replace(np.arange(25,100), "High")
smoker_deaths=pd.DataFrame({'total_deaths':covid_data.groupby('smokers')
).new_deaths.sum()})
smoker_deaths

covid_data_asia=covid_data.continent=='Asia'
plt.figure(figsize=(20,10))
# covid_data_cont.isnull=0
plt.title('Total deaths due to covid-19 of smokers ')

sns.barplot(x=smoker_deaths.index, y= smoker_deaths.total_deaths)
plt.legend
covid_data_asia=covid_data.continent=='Asia'
plt.figure(figsize=(20,10))
# covid_data_cont.isnull=0
plt.title('Spread of covid-19 & Vaccination')

sns.scatterplot(x=covid_data.people_vaccinated_per_thousand , y= covid_data.new_cases)
plt.legend()

covid_data_asia=covid_data.continent=='Asia'
plt.figure(figsize=(20,10))
# covid_data_cont.isnull=0
plt.title('Total deaths due to covid-19 in different Continents')

sns.scatterplot(x=covid_data.total_cases , y= covid_data.total_deaths)
#2%

avg=covid_data.new_deaths.mean()
avg_cas=covid_data.new_cases.mean()
covid_data.new_deaths=covid_data.new_deaths.fillna(avg)
covid_data.new_cases=covid_data.new_cases.fillna(avg_cas)
r=(covid_data.new_deaths/covid_data.new_cases)*100

covid_data.insert(2, "Mortality", r)
covid_data

covid_data_asia=covid_data.continent=='Asia'
plt.figure(figsize=(20,10))
# covid_data_cont.isnull=0

```

```
plt.title('Comparison of gdp to mortality rate')
avg=covid_data.total_deaths.mean()
avg_cas=covid_data.total_cases.mean()

rate=(covid_data.total_deaths/covid_data.total_cases)*100
print(rate)
sns.scatterplot(x=covid_data.gdp_per_capita , y= covid_data.Mortality)
#2%

covid_data_asia=covid_data.continent=='Asia'
plt.figure(figsize=(20,10))
#covid_data_cont.isnull=0
plt.title('Comparison of median age to mortality rate')

sns.scatterplot(x=covid_data.median_age , y= covid_data.Mortality)
#2%
```