

Covid vaccine analysis

Phase 4 : Coding part

Project Title : Covid vaccine analysis with
python programming

Items in the dataset:

- Countries
- Dates
- Vaccines
- Total Vaccinations

Desired data to find:

- Most commonly used vaccines in countries
- Average daily vaccination count in countries -Number of countries where vaccines are used
- Choropleth map of the most used vaccine

INPUT:

```
data=pd.DataFrame(columns=['Country', 'Vaccine',  
'Total_vaccine'])  
  
for country in df["location"].unique():  
  
    for vaccine in df["vaccine"].unique():  
  
        filtered_data = df[(df['location'] == country) & (df['vaccine']  
== vaccine)]  
  
        total_count = filtered_data['total_vaccinations'].max()  
  
        data = pd.concat([data, pd.DataFrame({'Country': [country],  
'Vaccine': [vaccine], 'Total_vaccine': [total_count]})],  
ignore_index=True)
```

SUB-INPUT:

```
data.head(10)
```

OUTPUT:

	Country	Vaccine	Total_vaccine
0	Argentina	Moderna	6507561
1	Argentina	Oxford/AstraZeneca	25977231
2	Argentina	Sinopharm/Beijing	28322602
3	Argentina	Sputnik V	20405678
4	Argentina	CanSino	610540
5	Argentina	Pfizer/BioNTech	14681054
6	Argentina	Johnson&Johnson	NaN
7	Argentina	Novavax	NaN
8	Argentina	Sinovac	NaN
9	Argentina	Covaxin	NaN

Most commonly used vaccines

INPUT:

```
data_2=pd.DataFrame(columns=['Country', 'Vaccine'])
data["Total_vaccine"] = pd.to_numeric(data["Total_vaccine"],
errors="coerce")
for country in data["Country"].unique():
new_data = data[data["Country"] == country]
max_vaccine = new_data.loc[new_data["Total_vaccine"].idxmax(), "Vaccine"]
data_2 = pd.concat([data_2, pd.DataFrame({'Country': [country], 'Vaccine': [max_vaccine]})], ignore_index=True)
```

SUB-INPUT:

```
data_2.head()
```

OUTPUT:

	Country	Vaccine
0	Argentina	Sinopharm/Beijing
1	Austria	Pfizer/BioNTech
2	Belgium	Pfizer/BioNTech
3	Bulgaria	Pfizer/BioNTech
4	Chile	Sinovac

INPUT:

```
data_2["Vaccine"].value_counts().plot(kind="bar",  
color=["Red", "Gray", "Gray", "Gray"])
```

OUTPUT:

<Axes: >



