In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:
```
filepath=(r"C:\Users\Jhuvvi Sree\OneDrive\Desktop\ds1.csv")
data=pd.read_csv(filepath)
print(data)
```

```
     s.no                      country names Total confirmed cases  \
0       1                              India           4,49,94,575
1       2                              china           9,92,92,081
2       3                              italy           2,58,97,801
3       4                                USA          10,34,36,829
4       5                              spain           1,39,14,811
..    ...                                ...                   ...
167   168                            Grenada                19,693
168   169  Saint vincent and the grenalidies                 9,631
169   170                            Eritrea                10,189
170   171                             Gambia                12,626
171   172                         Mozambique             2,33,417

    Recovered cases     Deaths Vaccinated people
0       4,44,61,208   5,31,913    2,20,67,37,729
1          3,79,053   1,21,490    3,51,58,72,818
2       2,55,93,568   1,90,868      15,03,17,622
3      10,55,25,088  11,27,152      66,88,82,018
4       1,37,62,417   1,21,760      11,28,37,406
..              ...        ...               ...
167          19,358        238            90,686
168           9,493        124            73,443
169          10,086        103              ----
170          12,189        372         14,44,492
171        2,28,805      2,243       3,49,48,793

[172 rows x 6 columns]
```

In [3]:
```
data = data.loc[:, ~data.columns.str.contains('^Unnamed')]
```

In [4]:
```
print(data)
```

```
          s.no                      country names Total confirmed cases  \
0            1                              India           4,49,94,575
1            2                              china           9,92,92,081
2            3                              italy           2,58,97,801
3            4                                USA          10,34,36,829
4            5                              spain           1,39,14,811
..         ...                                ...                   ...
167        168                            Grenada                19,693
168        169  Saint vincent and the grenalidies                 9,631
169        170                            Eritrea                10,189
170        171                             Gambia                12,626
171        172                         Mozambique             2,33,417

    Recovered cases     Deaths Vaccinated people
0       4,44,61,208   5,31,913    2,20,67,37,729
1          3,79,053   1,21,490    3,51,58,72,818
2       2,55,93,568   1,90,868      15,03,17,622
3      10,55,25,088  11,27,152      66,88,82,018
4       1,37,62,417   1,21,760      11,28,37,406
..              ...        ...               ...
167          19,358        238            90,686
168           9,493        124            73,443
169          10,086        103              ----
170          12,189        372         14,44,492
171        2,28,805      2,243       3,49,48,793

[172 rows x 6 columns]
```

In [5]:
```
data.head()
```

Out[5]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
0	1	India	4,49,94,575	4,44,61,208	5,31,913	2,20,67,37,729
1	2	china	9,92,92,081	3,79,053	1,21,490	3,51,58,72,818
2	3	italy	2,58,97,801	2,55,93,568	1,90,868	15,03,17,622
3	4	USA	10,34,36,829	10,55,25,088	11,27,152	66,88,82,018
4	5	spain	1,39,14,811	1,37,62,417	1,21,760	11,28,37,406

In [6]:
```
data.describe()
```

Out[6]:

	s.no
count	172.000000
mean	86.500000
std	49.796252
min	1.000000
25%	43.750000
50%	86.500000
75%	129.250000
max	172.000000

In [7]:
```
data.tail()
```

Out[7]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
167	168	Grenada	19,693	19,358	238	90,686
168	169	Saint vincent and the grenalidies	9,631	9,493	124	73,443
169	170	Eritrea	10,189	10,086	103	----
170	171	Gambia	12,626	12,189	372	14,44,492
171	172	Mozambique	2,33,417	2,28,805	2,243	3,49,48,793

In [8]:
```
data.isnull().sum()
```

Out[8]:
```
     s.no                0
country names            0
Total confirmed cases    0
Recovered cases          1
Deaths                   0
Vaccinated people        0
dtype: int64
```

In [9]:
```
data.columns
```

Out[9]:
```
Index(['     s.no', 'country names', 'Total confirmed cases',
       'Recovered cases', 'Deaths', 'Vaccinated people'],
      dtype='object')
```

In [10]:
```
data.columns
```

Out[10]:
```
Index(['     s.no', 'country names', 'Total confirmed cases',
       'Recovered cases', 'Deaths', 'Vaccinated people'],
      dtype='object')
```

In [11]:
```
data.size
```

Out[11]: 1032

In [12]:
```
data.ndim
```

Out[12]: 2

In [13]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 172 entries, 0 to 171
Data columns (total 6 columns):
 #   Column                 Non-Null Count  Dtype 
---  ------                 --------------  ----- 
 0        s.no              172 non-null    int64 
 1   country names          172 non-null    object
 2   Total confirmed cases  172 non-null    object
 3   Recovered cases        171 non-null    object
 4   Deaths                 172 non-null    object
 5   Vaccinated people      172 non-null    object
dtypes: int64(1), object(5)
memory usage: 8.2+ KB
```

In [14]:
```
data.index
```

Out[14]: RangeIndex(start=0, stop=172, step=1)

In [15]:
```
data.count
```

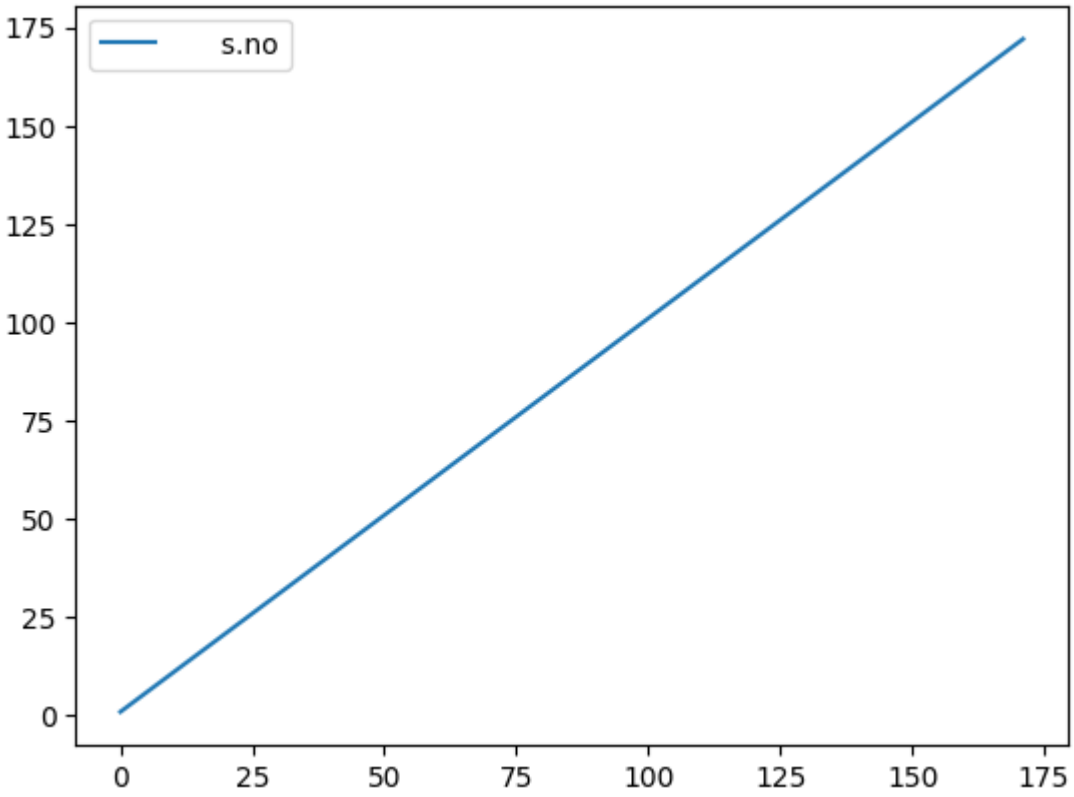
Out[15]:
```
<bound method DataFrame.count of          s.no                    countr
y names Total confirmed cases  \
0            1                             India          4,49,94,575
1            2                             china          9,92,92,081
2            3                             italy          2,58,97,801
3            4                               USA         10,34,36,829
4            5                             spain          1,39,14,811
..         ...                               ...                  ...
167        168                           Grenada               19,693
168        169  Saint vincent and the grenalidies                9,631
169        170                           Eritrea               10,189
170        171                            Gambia               12,626
171        172                         Mozambique            2,33,417

    Recovered cases     Deaths Vaccinated people
0       4,44,61,208   5,31,913    2,20,67,37,729
1          3,79,053   1,21,490    3,51,58,72,818
2       2,55,93,568   1,90,868      15,03,17,622
3      10,55,25,088  11,27,152      66,88,82,018
4       1,37,62,417   1,21,760      11,28,37,406
..              ...        ...               ...
167          19,358        238            90,686
168           9,493        124            73,443
169          10,086        103              ----
170          12,189        372          14,44,492
171       2,28,805      2,243       3,49,48,793

[172 rows x 6 columns]>
```

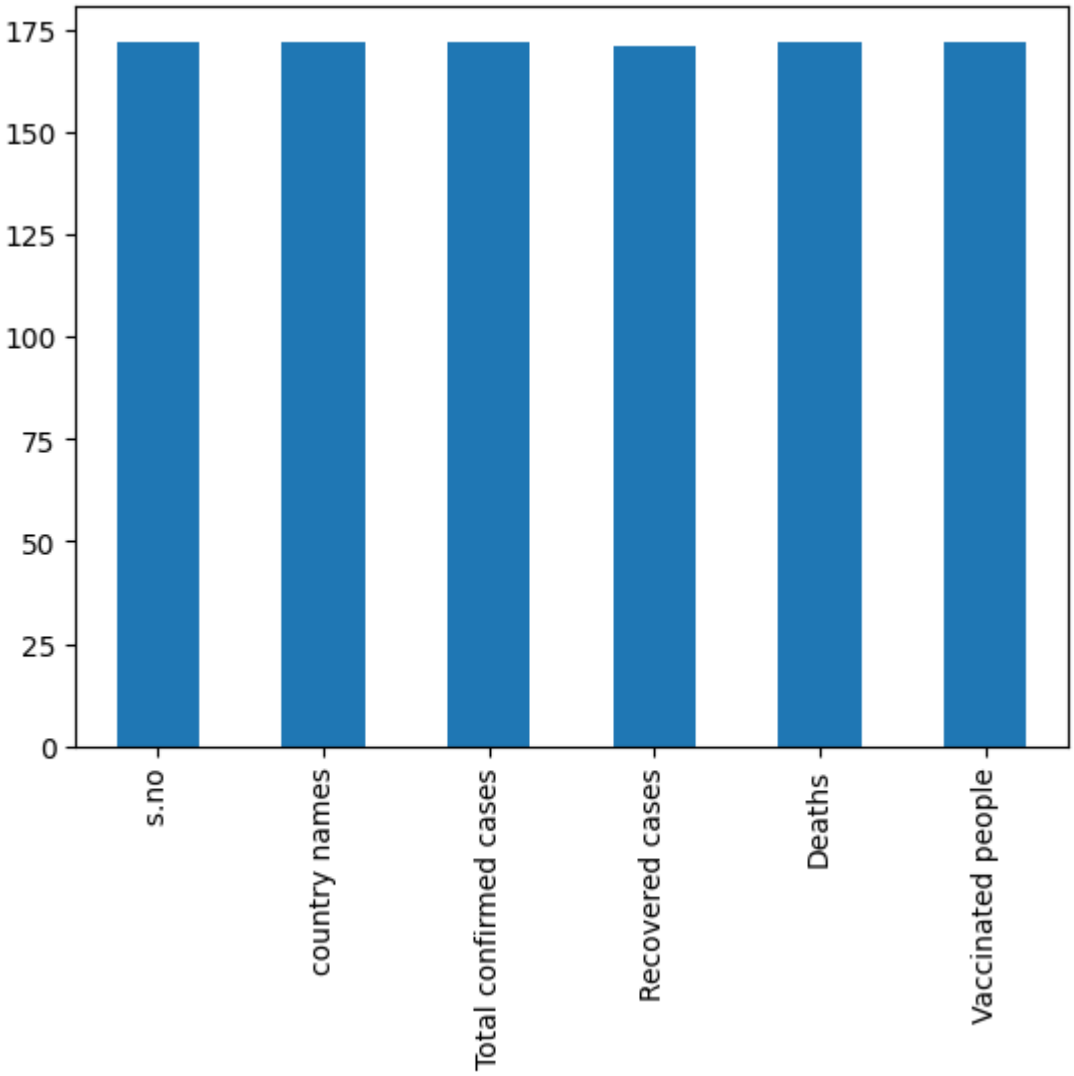
In [16]:
```
data.plot()
```

Out[16]:
```
<Axes: >
```



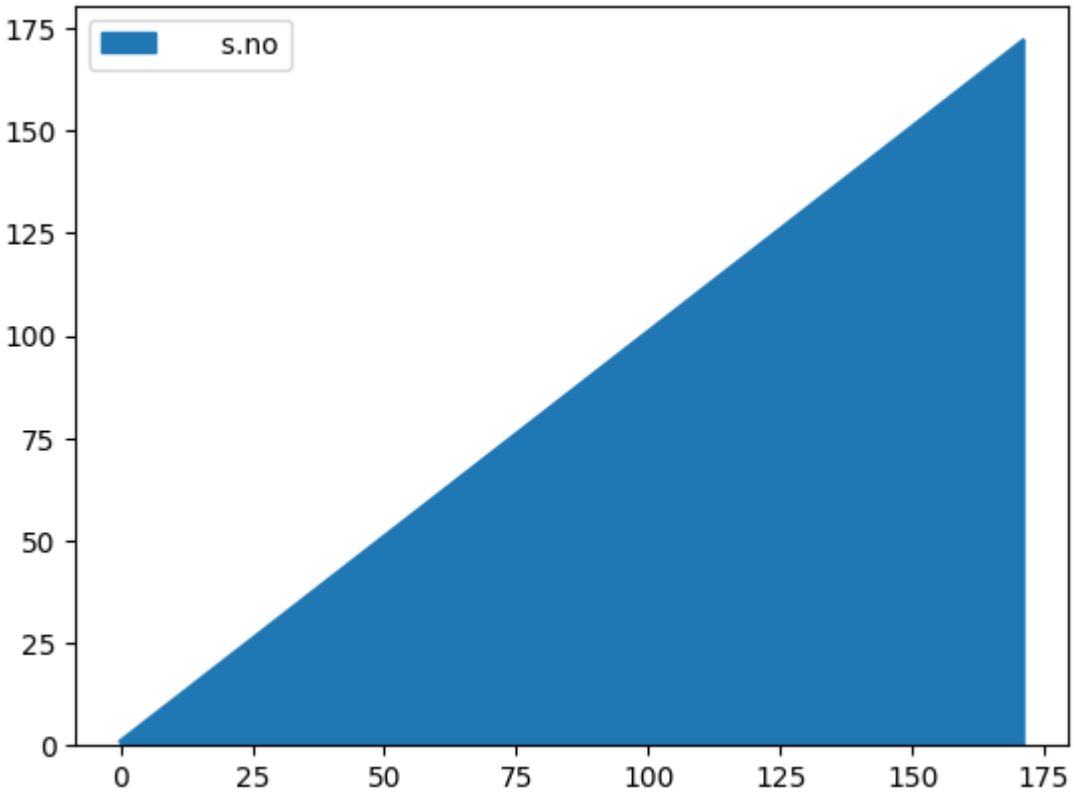
In [17]:
```
data.count().plot.bar()#barplot-all have equal entries
```

Out[17]: <Axes: >



In [18]:
```
data.plot.area()
```

Out[18]: <Axes: >



In [19]:
```
data
```

Out[19]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
0	1	India	4,49,94,575	4,44,61,208	5,31,913	2,20,67,37,729
1	2	china	9,92,92,081	3,79,053	1,21,490	3,51,58,72,818
2	3	italy	2,58,97,801	2,55,93,568	1,90,868	15,03,17,622
3	4	USA	10,34,36,829	10,55,25,088	11,27,152	66,88,82,018
4	5	spain	1,39,14,811	1,37,62,417	1,21,760	11,28,37,406
...	...	...	...	...	...	...
167	168	Grenada	19,693	19,358	238	90,686
168	169	Saint vincent and the grenalidies	9,631	9,493	124	73,443
169	170	Eritrea	10,189	10,086	103	----
170	171	Gambia	12,626	12,189	372	14,44,492
171	172	Mozambique	2,33,417	2,28,805	2,243	3,49,48,793

172 rows × 6 columns

In [20]:
```
x=data['Total confirmed cases']
x
```

Out[20]:
```
0       4,49,94,575
1       9,92,92,081
2       2,58,97,801
3      10,34,36,829
4       1,39,14,811
           ...
167          19,693
168           9,631
169          10,189
170          12,626
171        2,33,417
Name: Total confirmed cases, Length: 172, dtype: object
```

In [21]:
```
y=data['country names']
y
```

Out[21]:
```
0                                   India
1                                   china
2                                   italy
3                                     USA
4                                   spain
                      ...
167                              Grenada
168    Saint vincent and the grenalidies
169                              Eritrea
170                               Gambia
171                            Mozambique
Name: country names, Length: 172, dtype: object
```

In [22]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 172 entries, 0 to 171
Data columns (total 6 columns):
 #   Column                 Non-Null Count  Dtype 
---  ------                 --------------  ----- 
 0        s.no              172 non-null    int64 
 1   country names          172 non-null    object
 2   Total confirmed cases  172 non-null    object
 3   Recovered cases        171 non-null    object
 4   Deaths                 172 non-null    object
 5   Vaccinated people      172 non-null    object
dtypes: int64(1), object(5)
memory usage: 8.2+ KB
```

In [23]:
```
df=pd.DataFrame(data)
```

In [24]:
```
data
```

Out[24]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
0	1	India	4,49,94,575	4,44,61,208	5,31,913	2,20,67,37,729
1	2	china	9,92,92,081	3,79,053	1,21,490	3,51,58,72,818
2	3	italy	2,58,97,801	2,55,93,568	1,90,868	15,03,17,622
3	4	USA	10,34,36,829	10,55,25,088	11,27,152	66,88,82,018
4	5	spain	1,39,14,811	1,37,62,417	1,21,760	11,28,37,406
...	...	...	...	...	...	...
167	168	Grenada	19,693	19,358	238	90,686
168	169	Saint vincent and the grenalidies	9,631	9,493	124	73,443
169	170	Eritrea	10,189	10,086	103	----
170	171	Gambia	12,626	12,189	372	14,44,492
171	172	Mozambique	2,33,417	2,28,805	2,243	3,49,48,793

172 rows × 6 columns

In [25]:
```
df=pd.DataFrame(data)
```

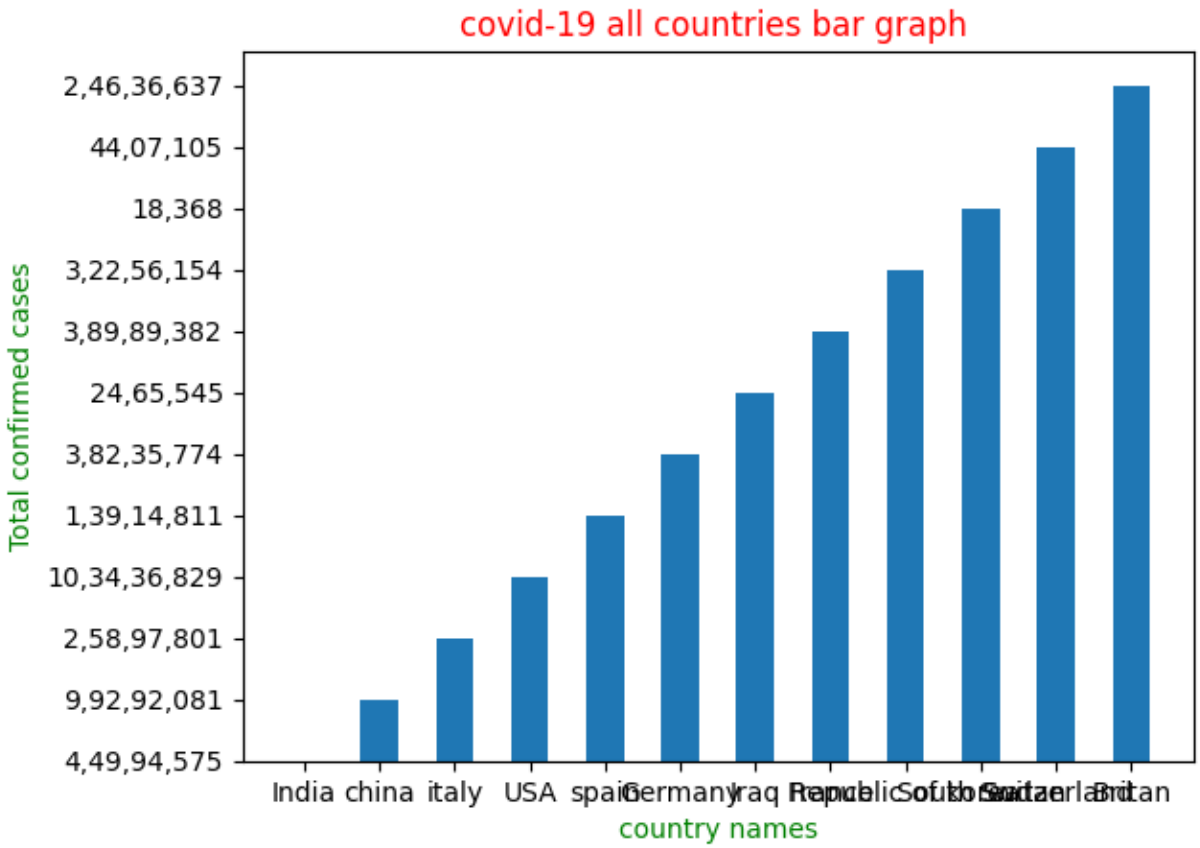
In [26]:
```
name = df['country names'].head(173)
num = df['Total confirmed cases'].head(173)
```

In [27]:
```
fig = plt.figure(figsize =(150,150))
```

```
<Figure size 15000x15000 with 0 Axes>
```

In [28]:
```
plt.bar(name[0:12], num[0:12],width=0.5)
plt.xlabel("country names",color="green")
plt.ylabel("Total confirmed cases",color="green")
plt.title("covid-19 all countries bar graph ",color="red")
```

Out[28]:
```
Text(0.5, 1.0, 'covid-19 all countries bar graph ')
```



In [29]:
```
data
```

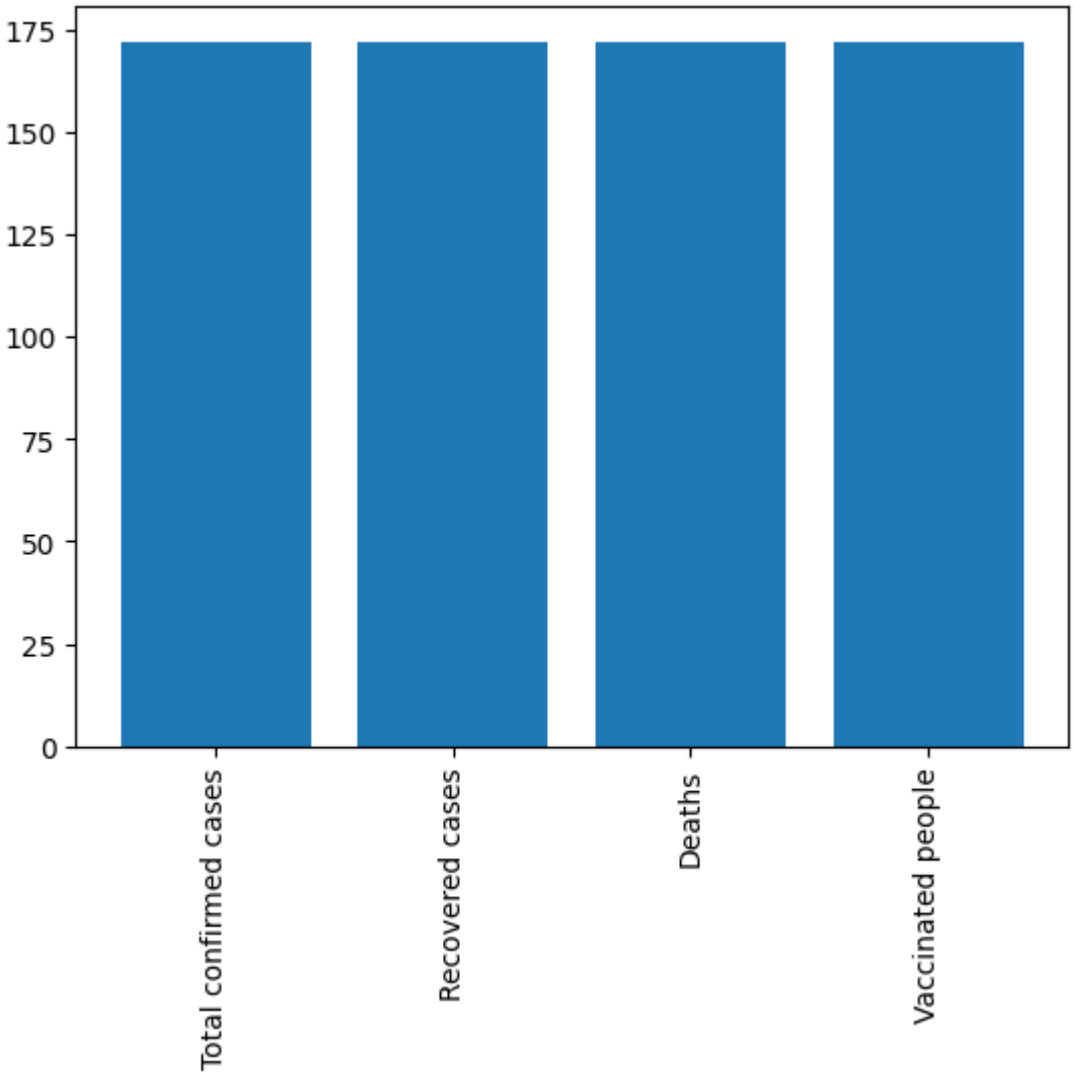
Out[29]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
0	1	India	4,49,94,575	4,44,61,208	5,31,913	2,20,67,37,729
1	2	china	9,92,92,081	3,79,053	1,21,490	3,51,58,72,818
2	3	italy	2,58,97,801	2,55,93,568	1,90,868	15,03,17,622
3	4	USA	10,34,36,829	10,55,25,088	11,27,152	66,88,82,018
4	5	spain	1,39,14,811	1,37,62,417	1,21,760	11,28,37,406
...	...	...	...	...	...	...
167	168	Grenada	19,693	19,358	238	90,686
168	169	Saint vincent and the grenalidies	9,631	9,493	124	73,443
169	170	Eritrea	10,189	10,086	103	----
170	171	Gambia	12,626	12,189	372	14,44,492
171	172	Mozambique	2,33,417	2,28,805	2,243	3,49,48,793

172 rows × 6 columns

In [30]:
```
heights = [172, 172, 172,172]
bars = ['Total confirmed cases', 'Recovered cases', 'Deaths','Vaccinated pec
y_pos = range(len(bars))
plt.bar(y_pos, heights)
# Rotation of the bars names
plt.xticks(y_pos, bars, rotation=90)
```

Out[30]:
```
([<matplotlib.axis.XTick at 0x2910a835e70>,
  <matplotlib.axis.XTick at 0x2910a835e40>,
  <matplotlib.axis.XTick at 0x2910a835a50>,
  <matplotlib.axis.XTick at 0x2910a4d67a0>],
 [Text(0, 0, 'Total confirmed cases'),
  Text(1, 0, 'Recovered cases'),
  Text(2, 0, 'Deaths'),
  Text(3, 0, 'Vaccinated people')])
```



In [31]:
```
data.sample(10)
```

Out[31]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
142	143	Saint Barthelemy	5,507	------	6	----
24	25	Israel	48,29,214	47,98,473	12,568	1,79,15,305
144	145	Cabo Verde	64,149	63,663	414	8,59,940
37	38	Iceland	2,09,191	-----	229	8,69,976
103	104	Democratic republic of congo	97,644	84,489	1,468	1,88,08,671
127	128	Uganda	1,70,775	1,00,431	3,632	2,64,06,936
125	126	Mongolia	10,10,034	9,41,511	2,179	56,68,144
41	42	Qatar	5,14,524	5,13,687	690	76,09,178
83	84	Burkina Faso	22,056	21,596	396	66,74,010
54	55	Lebanon	12,38,552	10,87,587	10,936	58,14,699

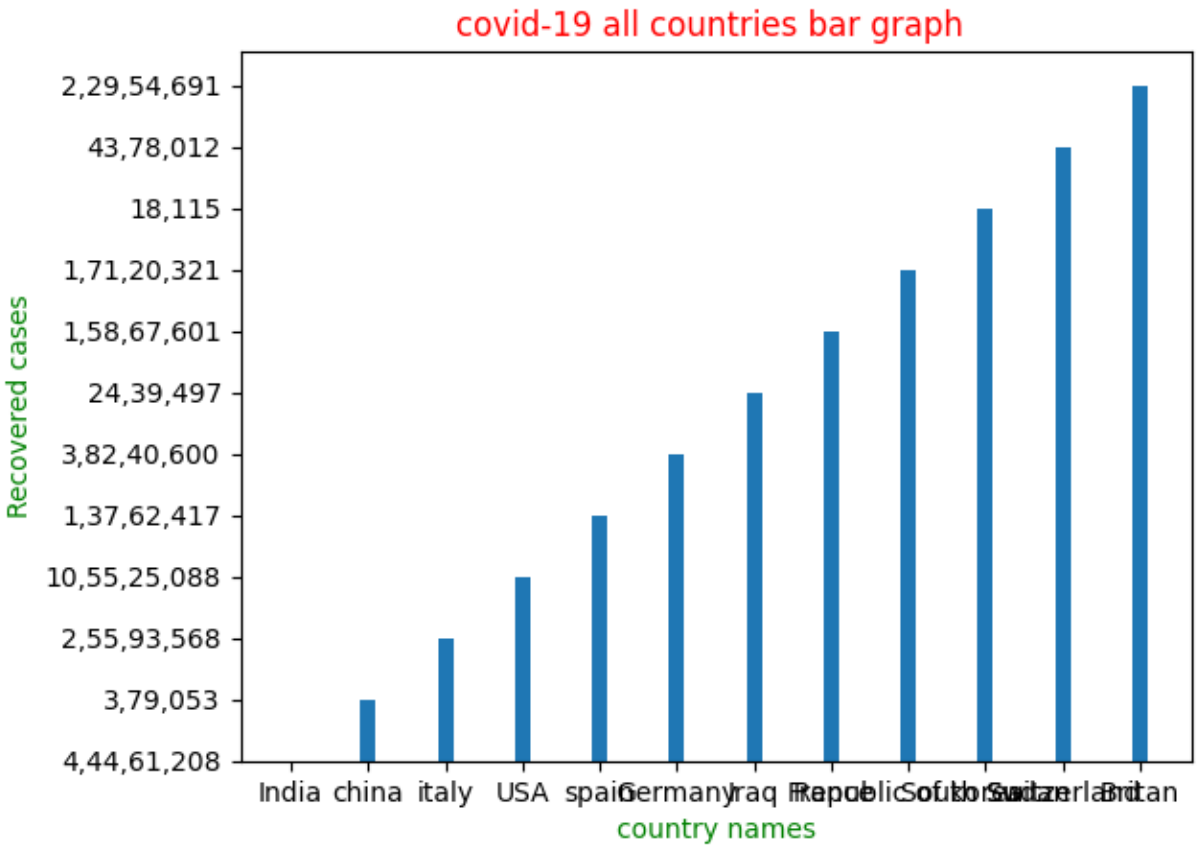
In [32]:
```
name1 = df['country names'].head(173)
num1 = df['Recovered cases'].head(173)
```

In [33]:
```
fig = plt.figure(figsize =(150,150))
```

```
<Figure size 15000x15000 with 0 Axes>
```

In [34]:
```
plt.bar(name1[0:12], num1[0:12],width=0.2)
plt.xlabel("country names",color="green")
plt.ylabel("Recovered cases",color="green")
plt.title("covid-19 all countries bar graph ",color="red")
```

Out[34]:
```
Text(0.5, 1.0, 'covid-19 all countries bar graph ')
```



In [35]:
```
data
```

Out[35]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
0	1	India	4,49,94,575	4,44,61,208	5,31,913	2,20,67,37,729
1	2	china	9,92,92,081	3,79,053	1,21,490	3,51,58,72,818
2	3	italy	2,58,97,801	2,55,93,568	1,90,868	15,03,17,622
3	4	USA	10,34,36,829	10,55,25,088	11,27,152	66,88,82,018
4	5	spain	1,39,14,811	1,37,62,417	1,21,760	11,28,37,406
...	...	...	...	...	...	...
167	168	Grenada	19,693	19,358	238	90,686
168	169	Saint vincent and the grenalidies	9,631	9,493	124	73,443
169	170	Eritrea	10,189	10,086	103	----
170	171	Gambia	12,626	12,189	372	14,44,492
171	172	Mozambique	2,33,417	2,28,805	2,243	3,49,48,793

172 rows × 6 columns

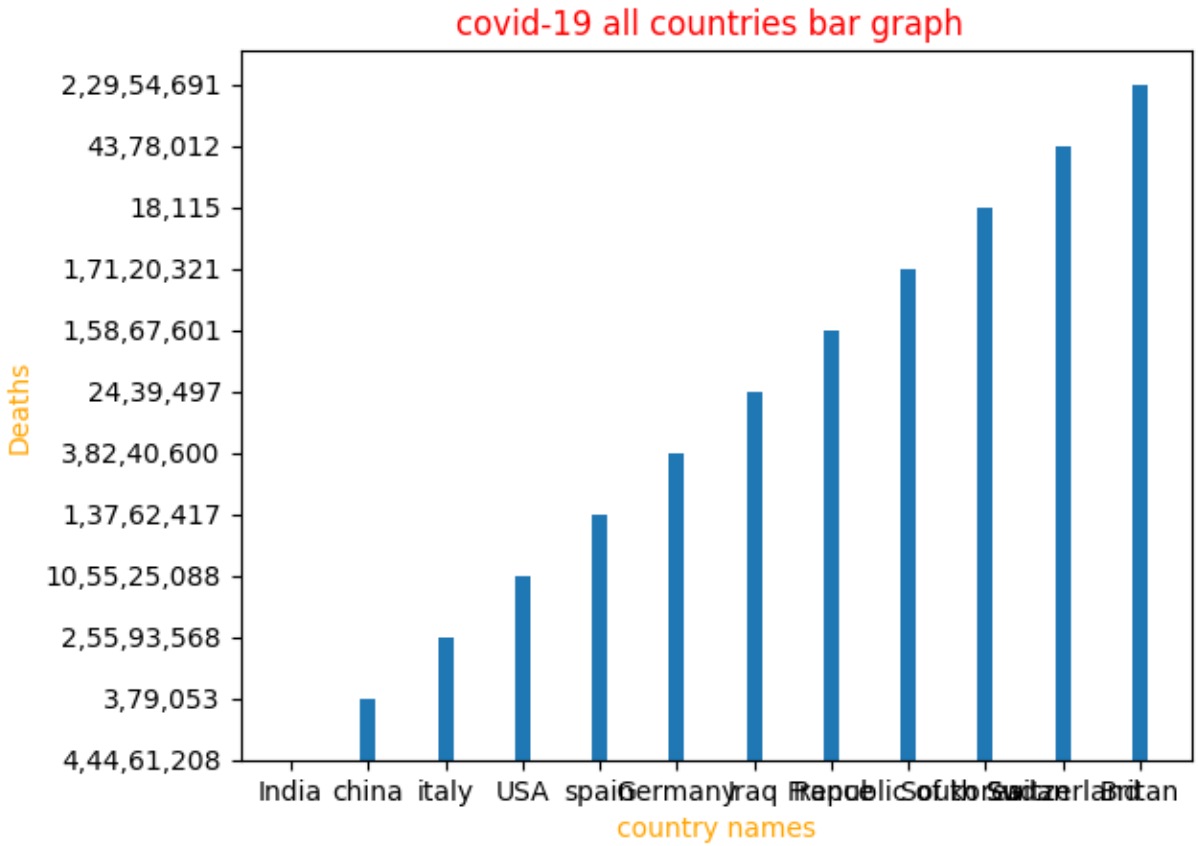
In [36]:
```
name2 = df['country names'].head(173)
num2 = df['Deaths'].head(173)
```

In [37]:
```
fig = plt.figure(figsize =(150,150))
```

```
<Figure size 15000x15000 with 0 Axes>
```

In [38]:
```
plt.bar(name1[0:12], num1[0:12],width=0.2)
plt.xlabel("country names",color="orange")
plt.ylabel("Deaths",color="orange")
plt.title("covid-19 all countries bar graph ",color="red")
```

Out[38]:
```
Text(0.5, 1.0, 'covid-19 all countries bar graph ')
```



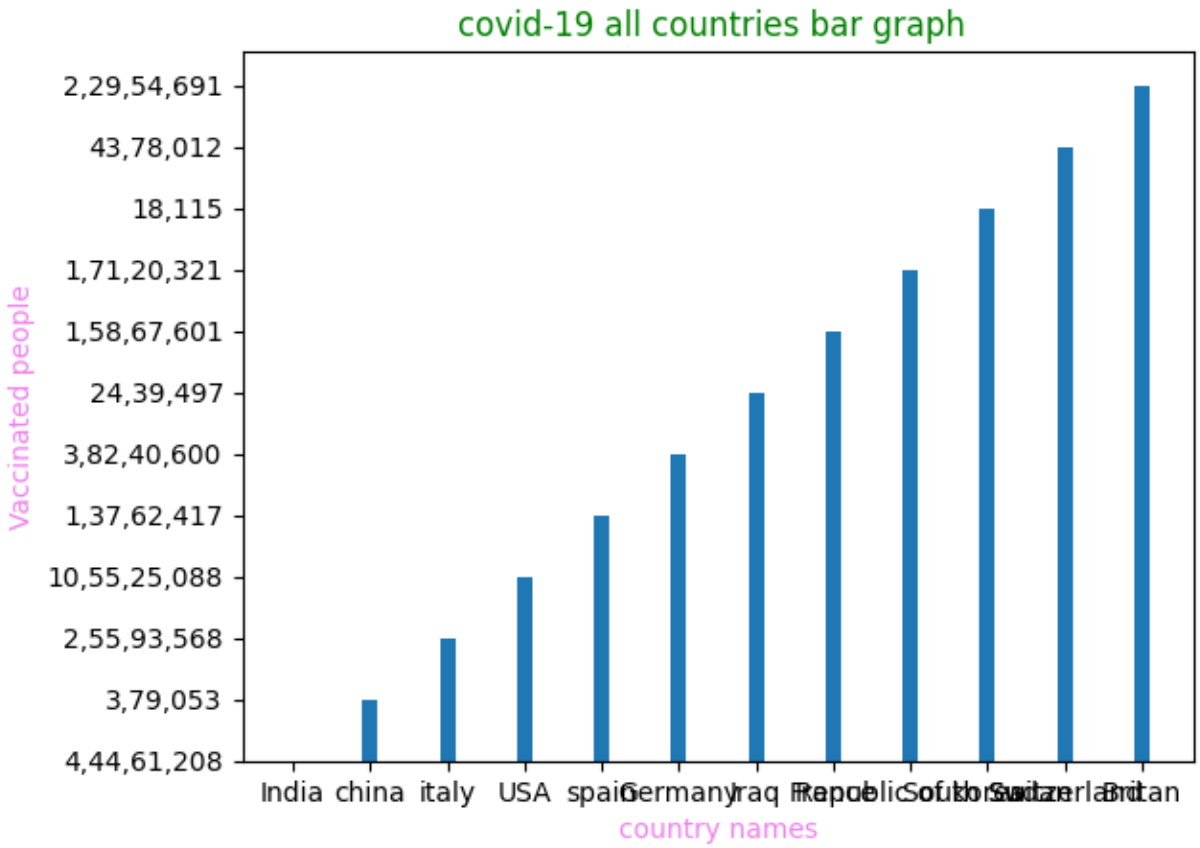
In [39]:
```
name3 = df['country names'].head(173)
num3 = df['Vaccinated people'].head(173)
```

In [40]:
```
fig = plt.figure(figsize =(150,150))
```

```
<Figure size 15000x15000 with 0 Axes>
```

In [41]:
```
plt.bar(name1[0:12], num1[0:12],width=0.2)
plt.xlabel("country names",color="violet")
plt.ylabel("Vaccinated people",color="violet")
plt.title("covid-19 all countries bar graph ",color="green")
```

Out[41]:
```
Text(0.5, 1.0, 'covid-19 all countries bar graph ')
```



In [42]:
```
data.isna().all()
```

Out[42]:
```
     s.no                 False
country names             False
Total confirmed cases     False
Recovered cases           False
Deaths                    False
Vaccinated people         False
dtype: bool
```

In [43]:
```
data
```

Out[43]:

	s.no	country names	Total confirmed cases	Recovered cases	Deaths	Vaccinated people
0	1	India	4,49,94,575	4,44,61,208	5,31,913	2,20,67,37,729
1	2	china	9,92,92,081	3,79,053	1,21,490	3,51,58,72,818
2	3	italy	2,58,97,801	2,55,93,568	1,90,868	15,03,17,622
3	4	USA	10,34,36,829	10,55,25,088	11,27,152	66,88,82,018
4	5	spain	1,39,14,811	1,37,62,417	1,21,760	11,28,37,406
...	...	...	...	...	...	...
167	168	Grenada	19,693	19,358	238	90,686
168	169	Saint vincent and the grenalidies	9,631	9,493	124	73,443
169	170	Eritrea	10,189	10,086	103	----
170	171	Gambia	12,626	12,189	372	14,44,492
171	172	Mozambique	2,33,417	2,28,805	2,243	3,49,48,793

172 rows × 6 columns