NAME - Kartikay Agrawal
REG NO. - 2148064

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [4]:

```
df = pd.read_csv('/content/sample_data/country_vaccinations (1).csv')
```

In [5]:

```
df.head()
```

Out[5]:

	country	iso_code	date	total_vaccinations	people_vaccinated	people_fully_vaccinated	d
0	Afghanistan	AFG	2021-02-22	0.0	0.0	NaN	
1	Afghanistan	AFG	2021-02-23	NaN	NaN	NaN	
2	Afghanistan	AFG	2021-02-24	NaN	NaN	NaN	
3	Afghanistan	AFG	2021-02-25	NaN	NaN	NaN	
4	Afghanistan	AFG	2021-02-26	NaN	NaN	NaN	

In [6]:

```
df.info()
```

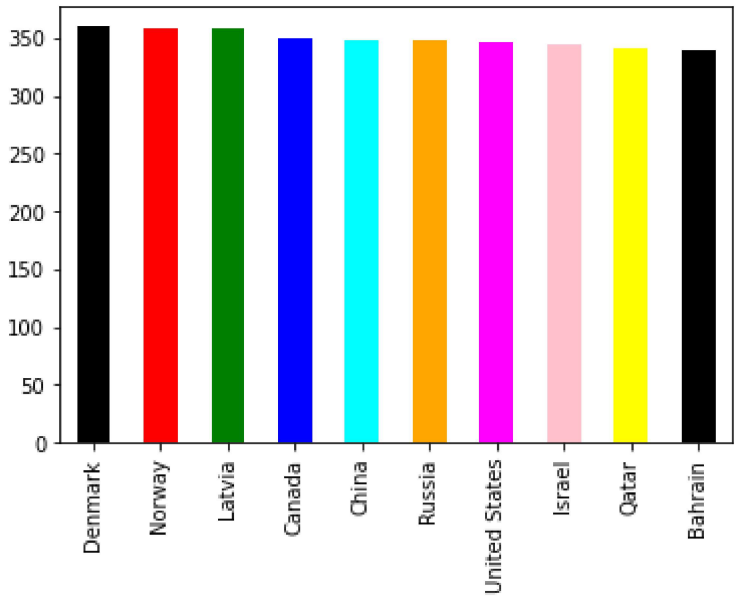
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 60269 entries, 0 to 60268
Data columns (total 15 columns):
 #   Column                                Non-Null Count  Dtype  
---  ------                                --------------  -----  
 0   country                               60269 non-null  object 
 1   iso_code                              60269 non-null  object 
 2   date                                  60269 non-null  object 
 3   total_vaccinations                    32315 non-null  float64
 4   people_vaccinated                     30739 non-null  float64
 5   people_fully_vaccinated               27943 non-null  float64
 6   daily_vaccinations_raw                26398 non-null  float64
 7   daily_vaccinations                    59936 non-null  float64
 8   total_vaccinations_per_hundred        32315 non-null  float64
 9   people_vaccinated_per_hundred         30739 non-null  float64
 10  people_fully_vaccinated_per_hundred   27943 non-null  float64
 11  daily_vaccinations_per_million        59936 non-null  float64
 12  vaccines                              60269 non-null  object 
 13  source_name                           60269 non-null  object 
 14  source_website                        60269 non-null  object 
dtypes: float64(9), object(6)
memory usage: 6.9+ MB
```

In [22]:

```
df['country'].value_counts().head(10).plot.bar(color=['black', 'red', 'green', 'blue',
```

Out[22]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcf23de0110>
```

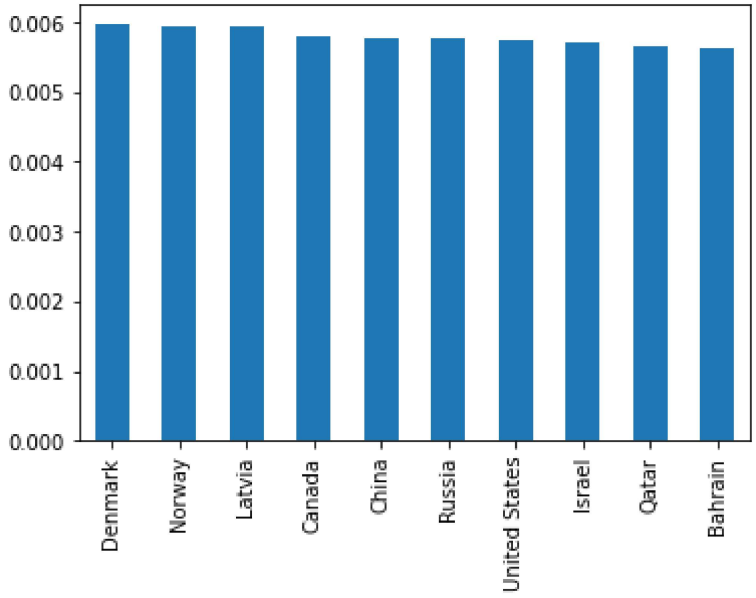


In [23]:

```
(df['country'].value_counts().head(10) / len(df)).plot.bar()
```

Out[23]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcf23e3d250>
```

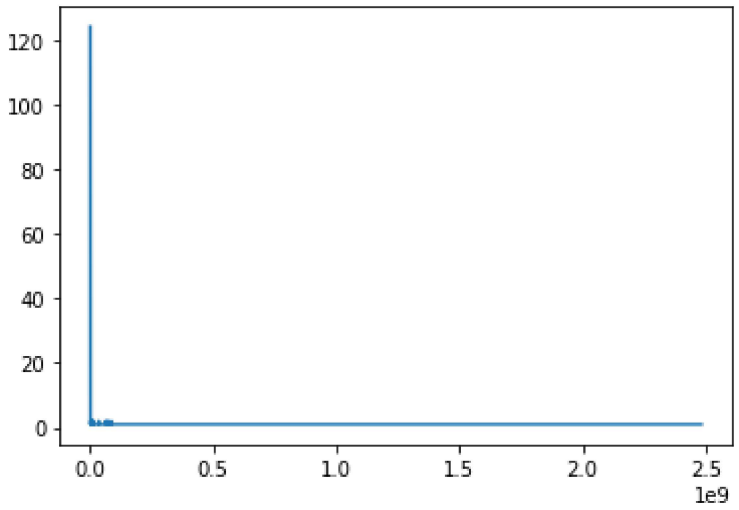


In [24]:

```
df['total_vaccinations'].value_counts().sort_index().plot.line()
```

Out[24]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcf23cf0090>
```

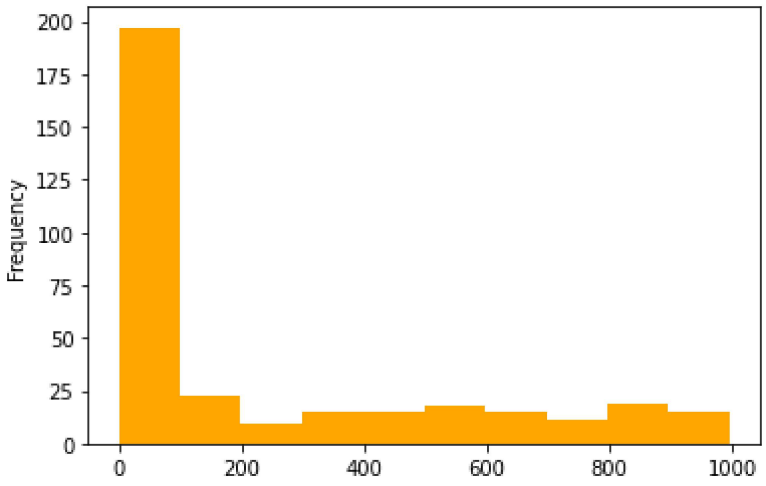


In [26]:

```
df[df['total_vaccinations'] < 1000]['total_vaccinations'].plot.hist(color=['orange'])
```

Out[26]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcf23c4d3d0>
```



In [27]:

```
pd.pivot_table(df,index=["country"])
```

Out[27]:

country	daily_vaccinations	daily_vaccinations_per_million	daily_vaccinations_raw	people_fully
Afghanistan	16869.269373	423.450185	3437.000000	4.
Albania	6456.470219	2247.300940	8311.044304	5.
Algeria	39893.531987	893.962963	11778.000000	4.
Andorra	372.580645	4816.519713	2401.000000	2.
Angola	32478.626866	957.111940	NaN	9.
...	...	...	...	
Wales	17034.715625	5373.725000	17353.006369	1.
Wallis and Futuna	46.770492	4215.840164	NaN	4.
Yemen	3822.892857	125.372449	NaN	1
Zambia	4619.396476	244.110132	6955.678161	2.
Zimbabwe	23009.263345	1524.569395	24531.350427	1

223 rows × 9 columns

In [28]:

```
df.query('daily_vaccinations_raw == `daily_vaccinations`')
```

Out[28]:

	country	iso_code	date	total_vaccinations	people_vaccinated	people_fully_vaccinat
593	Algeria	DZA	2021-01-30	30.0	30.0	Na
2772	Aruba	ABW	2021-10-08	157355.0	81761.0	7559
2824	Australia	AUS	2021-02-22	2789.0	2789.0	Na
3104	Austria	AUT	2020-12-28	2092.0	2092.0	Na
3755	Bahamas	BHS	2021-03-14	110.0	100.0	Na
...	...	...	...	...	...	
58735	Vietnam	VNM	2021-03-08	377.0	377.0	Na
58997	Wales	OWID_WLS	2021-01-11	91336.0	91239.0	9
59247	Wales	OWID_WLS	2021-09-18	4582013.0	2368230.0	221378
59760	Zambia	ZMB	2021-04-15	106.0	106.0	Na
60019	Zimbabwe	ZWE	2021-03-22	44120.0	43840.0	28

126 rows × 15 columns

In [29]:

```
df['total_vaccinations'].agg(['sum', 'min'])
```

Out[29]:

```
sum    9.405338e+11
min    0.000000e+00
Name: total_vaccinations, dtype: float64
```