

```
In [ ]: !pip install pandas
!pip install numpy
!pip install plotly
```

```
In [5]: !pip install plotly
```

```
Requirement already satisfied: plotly in c:\users\yugmee\appdata\local\programs\python\python39\lib\site-packages (5.1.0)
Requirement already satisfied: six in c:\users\yugmee\appdata\local\programs\python\python39\lib\site-packages (from plotly) (1.15.0)
Requirement already satisfied: tenacity>=6.2.0 in c:\users\yugmee\appdata\local\programs\python\python39\lib\site-packages (from plotly) (7.0.0)
```

```
In [6]: import pandas as pd
import numpy as pn
import plotly.express as px
```

```
-----
ModuleNotFoundError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_24968\1514410709.py in <module>
      1 import pandas as pd
----> 2 import numpy as pn
      3 import plotly.express as px

ModuleNotFoundError: No module named 'numpy'
```

```
In [ ]: px
```

```
In [9]: import numpy as np
```

```
In [12]: import pandas as pd
import numpy as pn
import plotly.express as px
```

In [13]: ▶ px

Out[13]: <module 'plotly.express' from 'c:\\users\\yugmee\\appdata\\local\\programs\\python\\python39\\lib\\site-packages\\plotly\\express__init__.py'>

In [14]: ▶ `!curl -s /raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv`
`head_csv(base_url + 'time_series_covid19_confirmed_global.csv')`

562

563 # XXX probably also want an abstract factory that knows when it makes

c:\\users\\yugmee\\appdata\\local\\programs\\python\\python39\\lib\\urllib\\request.py in _call_chain(self, chain, kind, meth_name, *args)

492 for handler in handlers:

493 func = getattr(handler, meth_name)

--> 494 result = func(*args)

495 if result is not None:

496 return result

c:\\users\\yugmee\\appdata\\local\\programs\\python\\python39\\lib\\urllib\\request.py in http_error_default(self, req, fp, code, msg, hdrs)

639 class HTTPDefaultErrorHandler(BaseHandler):

640 def http_error_default(self, req, fp, code, msg, hdrs):

--> 641 raise HTTPError(req.full_url, code, msg, hdrs, fp)

642

643 class HTTPRedirectHandler(BaseHandler):

HTTPError: HTTP Error 404: Not Found

In []: ▶

```
In [15]: ► base_url='https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_df=pd.read_csv(base_url+ 'time_series_covid19_confirmed_global.csv')
```

```
561         return self._call_chain(*args)
562
563 # XXX probably also want an abstract factory that knows when it makes

c:\users\yugmee\appdata\local\programs\python\python39\lib\urllib\request.py in _call_chain(self, chain, kind, meth_name, *args)
492     for handler in handlers:
493         func = getattr(handler, meth_name)
--> 494         result = func(*args)
495         if result is not None:
496             return result

c:\users\yugmee\appdata\local\programs\python\python39\lib\urllib\request.py in http_error_default(self, req, fp, code, msg, hdrs)
639 class HTTPDefaultErrorHandler(BaseHandler):
640     def http_error_default(self, req, fp, code, msg, hdrs):
--> 641         raise HTTPError(req.full_url, code, msg, hdrs, fp)
642
643 class HTTPRedirectHandler(BaseHandler):
```

HTTPError: HTTP Error 404: Not Found

```
In [16]: ► confirmed_df
```

```
-----
NameError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_24968\2052993400.py in <module>
----> 1 confirmed_df

NameError: name 'confirmed_df' is not defined
```

```
In [18]: ► base_url='https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_df=pd.read_csv(base_url+ 'time_series_covid19_confirmed_global.csv')
```

```
In [19]: confirmed_df=pd.read_csv(base_url + 'time_series_covid19_confirmed_global.csv')
```

```
--> 610         return _read(filepath_or_buffer, kwds)
      611
      612
```

```
c:\users\yugmee\appdata\local\programs\python\python39\lib\site-packages\pandas\io\parsers.py in _read(filepath_or_buffer, kwds)
```

```
      460
      461     # Create the parser.
--> 462     parser = TextFileReader(filepath_or_buffer, **kwds)
      463
      464     if chunksize or iterator:
```

```
c:\users\yugmee\appdata\local\programs\python\python39\lib\site-packages\pandas\io\parsers.py in __init__(self, f, engine, **kwds)
```

```
      817         self.options["has_index_names"] = kwds["has_index_names"]
      818
--> 819         self._engine = self._make_engine(self.engine)
      820
      821     def close(self):
```

```
c:\users\yugmee\appdata\local\programs\python\python39\lib\site-packages\pandas\io\parsers.py in _make_engine(self, engine)
```

```
     1048         )
     1049         ..
     1050         ..
```

```
In [22]: confirmed_df=pd.read_csv(base_url)
```

```
In [23]: new_url=base_url+'time_series_covid19_confirmed_global.csv'
```

In [24]:

confirmed_df

Out[24]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	...	6/23/21	6/24/21	6/25/21	6/26/21	6/27/21	6/28/21	6/29/21	6/30/21
0	NaN	Afghanistan	33.939110	67.709953	0	0	0	0	0	0	...	107957	109532	111592	111592	111592	115751	117158	118
1	NaN	Albania	41.153300	20.168300	0	0	0	0	0	0	...	132497	132499	132506	132509	132512	132513	132514	132
2	NaN	Algeria	28.033900	1.659600	0	0	0	0	0	0	...	137049	137403	137772	138113	138465	138840	139229	139
3	NaN	Andorra	42.506300	1.521800	0	0	0	0	0	0	...	13873	13877	13882	13882	13882	13882	13900	13
4	NaN	Angola	-11.202700	17.873900	0	0	0	0	0	0	...	38002	38091	38371	38528	38556	38613	38682	38
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
274	NaN	Vietnam	14.058324	108.277199	0	2	2	2	2	2	...	13989	14323	14537	15325	15740	16136	16507	17
275	NaN	West Bank and Gaza	31.952200	35.233200	0	0	0	0	0	0	...	313015	313015	313520	313520	313721	313851	313851	314
276	NaN	Yemen	15.552727	48.516388	0	0	0	0	0	0	...	6898	6905	6905	6906	6908	6909	6918	6
277	NaN	Zambia	-13.133897	27.849332	0	0	0	0	0	0	...	137026	140620	143215	146031	148568	149661	152056	154
278	NaN	Zimbabwe	-19.015438	29.154857	0	0	0	0	0	0	...	43480	44306	45217	46018	46442	47284	48533	49

279 rows × 532 columns

In [25]:

confirmed_df=confirmed_df.drop(columns=['Lat','Long','Province/State'])

In [26]:

confirmed_df

Out[26]:

	Country/Region	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	1/28/20	1/29/20	1/30/20	...	6/23/21	6/24/21	6/25/21	6/26/21	6/27/21	6/28/21	6/29/21	6/30/21	7/1/21
0	Afghanistan	0	0	0	0	0	0	0	0	0	...	107957	109532	111592	111592	111592	115751	117158	118659	120216
1	Albania	0	0	0	0	0	0	0	0	0	...	132497	132499	132506	132509	132512	132513	132514	132521	132523
2	Algeria	0	0	0	0	0	0	0	0	0	...	137049	137403	137772	138113	138465	138840	139229	139626	140075
3	Andorra	0	0	0	0	0	0	0	0	0	...	13873	13877	13882	13882	13882	13882	13900	13911	13918
4	Angola	0	0	0	0	0	0	0	0	0	...	38002	38091	38371	38528	38556	38613	38682	38849	38965
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
274	Vietnam	0	2	2	2	2	2	2	2	2	...	13989	14323	14537	15325	15740	16136	16507	17052	17727
275	West Bank and Gaza	0	0	0	0	0	0	0	0	0	...	313015	313015	313520	313520	313721	313851	313851	314167	314288
276	Yemen	0	0	0	0	0	0	0	0	0	...	6898	6905	6905	6906	6908	6909	6918	6920	6923
277	Zambia	0	0	0	0	0	0	0	0	0	...	137026	140620	143215	146031	148568	149661	152056	154948	157832
278	Zimbabwe	0	0	0	0	0	0	0	0	0	...	43480	44306	45217	46018	46442	47284	48533	49864	51221

279 rows × 529 columns

In [27]:

confirmed_df=confirmed_df.groupby(by='Country/Region').aggregate(np.sum).T

In [28]:

confirmed_df

Out[28]:

Country/Region	Afghanistan	Albania	Algeria	Andorra	Angola	Antigua and Barbuda	Argentina	Armenia	Australia	Austria	...	United Kingdom	Uruguay	Uzbekistan	Vanuatu	Venezuela	Vietnam
1/22/20	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
1/23/20	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	2
1/24/20	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	2
1/25/20	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	2
1/26/20	0	0	0	0	0	0	0	0	4	0	...	0	0	0	0	0	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	..
6/28/21	115751	132513	138840	13882	38613	1263	4423636	224851	30560	650324	...	4771367	366915	109692	4	269635	16136
6/29/21	117158	132514	139229	13900	38682	1263	4447701	224967	30602	650353	...	4791628	368178	110190	4	270654	16507
6/30/21	118659	132521	139626	13911	38849	1263	4470374	225095	30643	650412	...	4817298	369350	110677	4	272712	17052
7/1/21	120216	132523	140075	13918	38965	1264	4491551	225221	30684	650474	...	4844944	370600	111153	4	274024	17727
7/2/21	120216	132526	140550	13918	39089	1264	4512439	225339	30734	650569	...	4871807	371417	111648	4	275205	18121

528 rows × 195 columns

In [29]:

confirmed_df.index.name='Date'
confirmed_df=confirmed_df.reset_index()

In [30]:

▶ confirmed_df.tail()

Out[30]:

Country/Region	Date	Afghanistan	Albania	Algeria	Andorra	Angola	Antigua and Barbuda	Argentina	Armenia	Australia	...	United Kingdom	Uruguay	Uzbekistan	Vanuatu	Venezuela	Vietnam
523	6/28/21	115751	132513	138840	13882	38613	1263	4423636	224851	30560	...	4771367	366915	109692	4	269635	16136
524	6/29/21	117158	132514	139229	13900	38682	1263	4447701	224967	30602	...	4791628	368178	110190	4	270654	16507
525	6/30/21	118659	132521	139626	13911	38849	1263	4470374	225095	30643	...	4817298	369350	110677	4	272712	17052
526	7/1/21	120216	132523	140075	13918	38965	1264	4491551	225221	30684	...	4844944	370600	111153	4	274024	17727
527	7/2/21	120216	132526	140550	13918	39089	1264	4512439	225339	30734	...	4871807	371417	111648	4	275205	18121

5 rows × 196 columns

In [31]:

▶ confirmed_melt_df=confirmed_df.melt(id_vars='Date').copy()

In [32]:

▶ confirmed_melt_df

Out[32]:

	Date	Country/Region	value
0	1/22/20	Afghanistan	0
1	1/23/20	Afghanistan	0
2	1/24/20	Afghanistan	0
3	1/25/20	Afghanistan	0
4	1/26/20	Afghanistan	0
...	...	...	...
102955	6/28/21	Zimbabwe	47284
102956	6/29/21	Zimbabwe	48533
102957	6/30/21	Zimbabwe	49864
102958	7/1/21	Zimbabwe	51221
102959	7/2/21	Zimbabwe	52663

102960 rows × 3 columns

In [33]:

▶ confirmed_melt_df.rename(columns={'value': 'Confirmed'}, inplace=True)

```
In [34]: confirmed_melt_df.head()
```

Out[34]:

	Date	Country/Region	Confirmed
0	1/22/20	Afghanistan	0
1	1/23/20	Afghanistan	0
2	1/24/20	Afghanistan	0
3	1/25/20	Afghanistan	0
4	1/26/20	Afghanistan	0

```
In [35]: max_date=confirmed_melt_df['Date'].max()  
max_date
```

Out[35]: '9/9/20'

```
In [37]: confirmed_melt_df['Date']=pd.to_datetime(confirmed_melt_df['Date'])
```

```
In [38]: confirmed_melt_df.head()
```

Out[38]:

	Date	Country/Region	Confirmed
0	2020-01-22	Afghanistan	0
1	2020-01-23	Afghanistan	0
2	2020-01-24	Afghanistan	0
3	2020-01-25	Afghanistan	0
4	2020-01-26	Afghanistan	0

```
In [39]: confirmed_melt_df['Date']=confirmed_melt_df['Date'].dt.strftime('%m/%d/%Y')
```

```
In [40]: max_date=confirmed_melt_df['Date'].max()
max_date
```

Out[40]: '12/31/2020'

```
In [41]: #visulaization
total_confirmed_df=confirmed_melt_df[confirmed_melt_df['Date']==max_date]
total_confirmed_df
```

Out[41]:

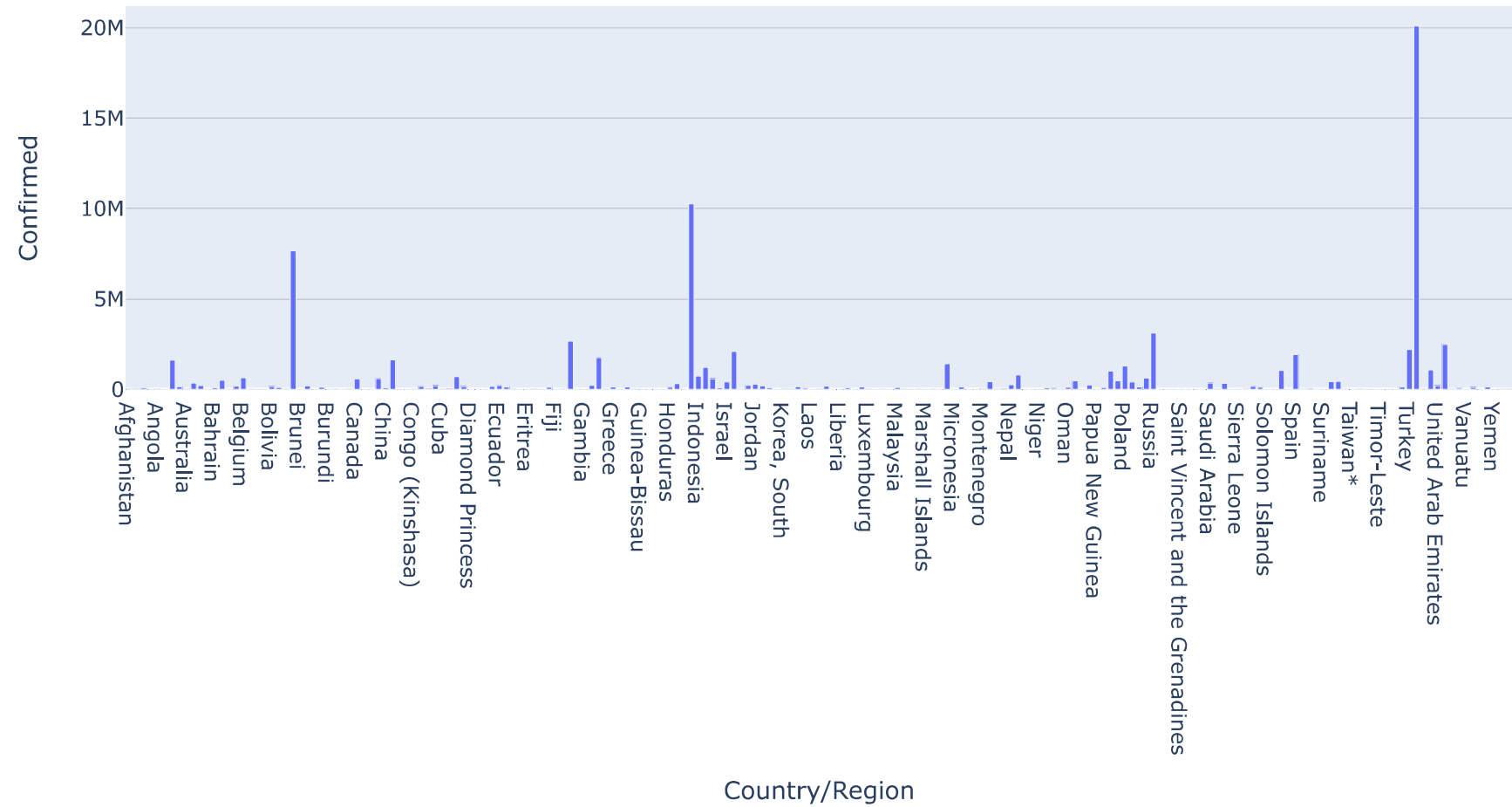
	Date	Country/Region	Confirmed
344	12/31/2020	Afghanistan	51526
872	12/31/2020	Albania	58316
1400	12/31/2020	Algeria	99610
1928	12/31/2020	Andorra	8049
2456	12/31/2020	Angola	17553
...	...	...	...
100664	12/31/2020	Vietnam	1465
101192	12/31/2020	West Bank and Gaza	138004
101720	12/31/2020	Yemen	2099
102248	12/31/2020	Zambia	20725
102776	12/31/2020	Zimbabwe	13867

195 rows × 3 columns

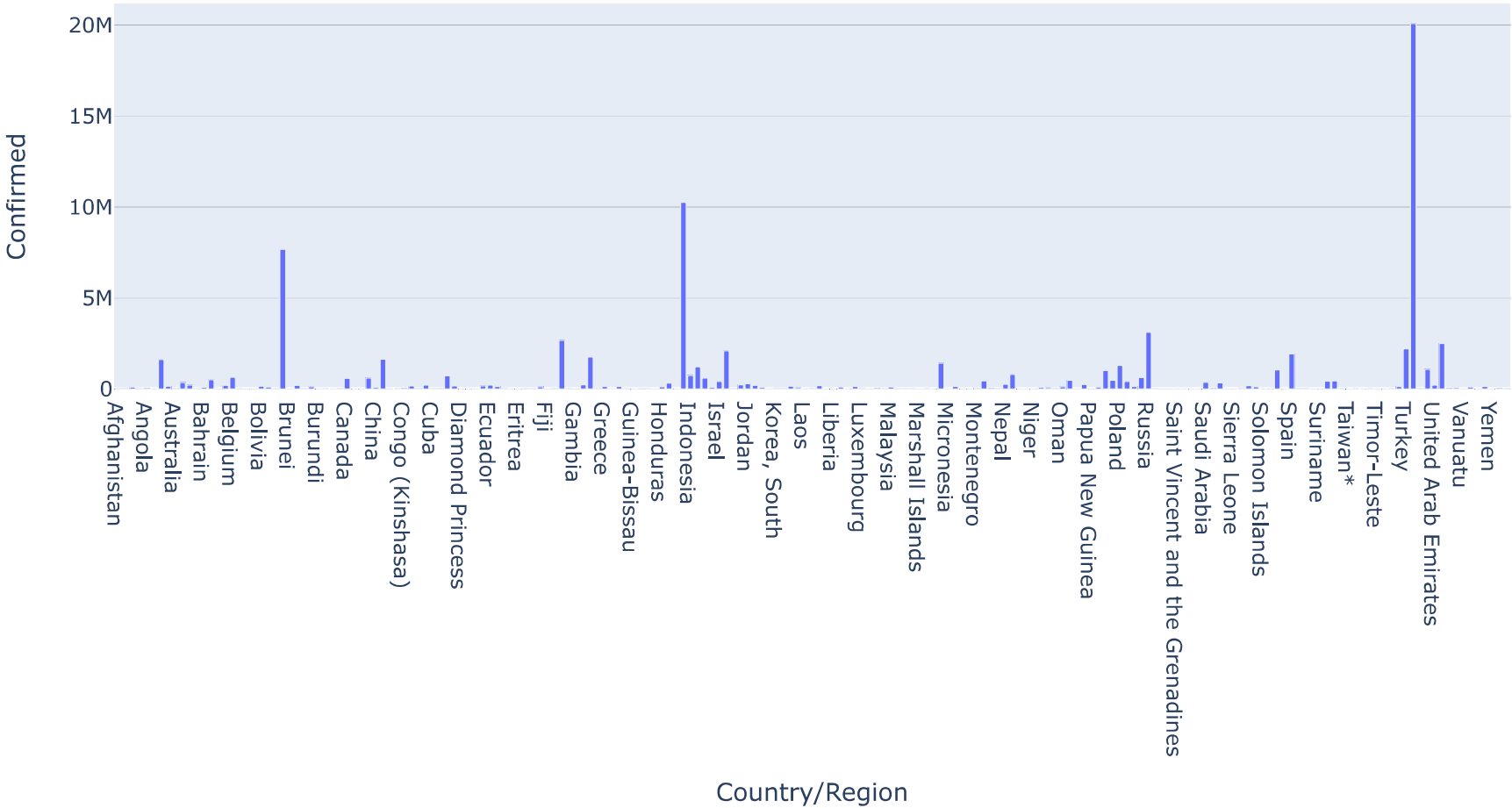
```
In [42]: total_confirmed= total_confirmed_df['Confirmed'].sum()
total_confirmed
```

Out[42]: 83560482

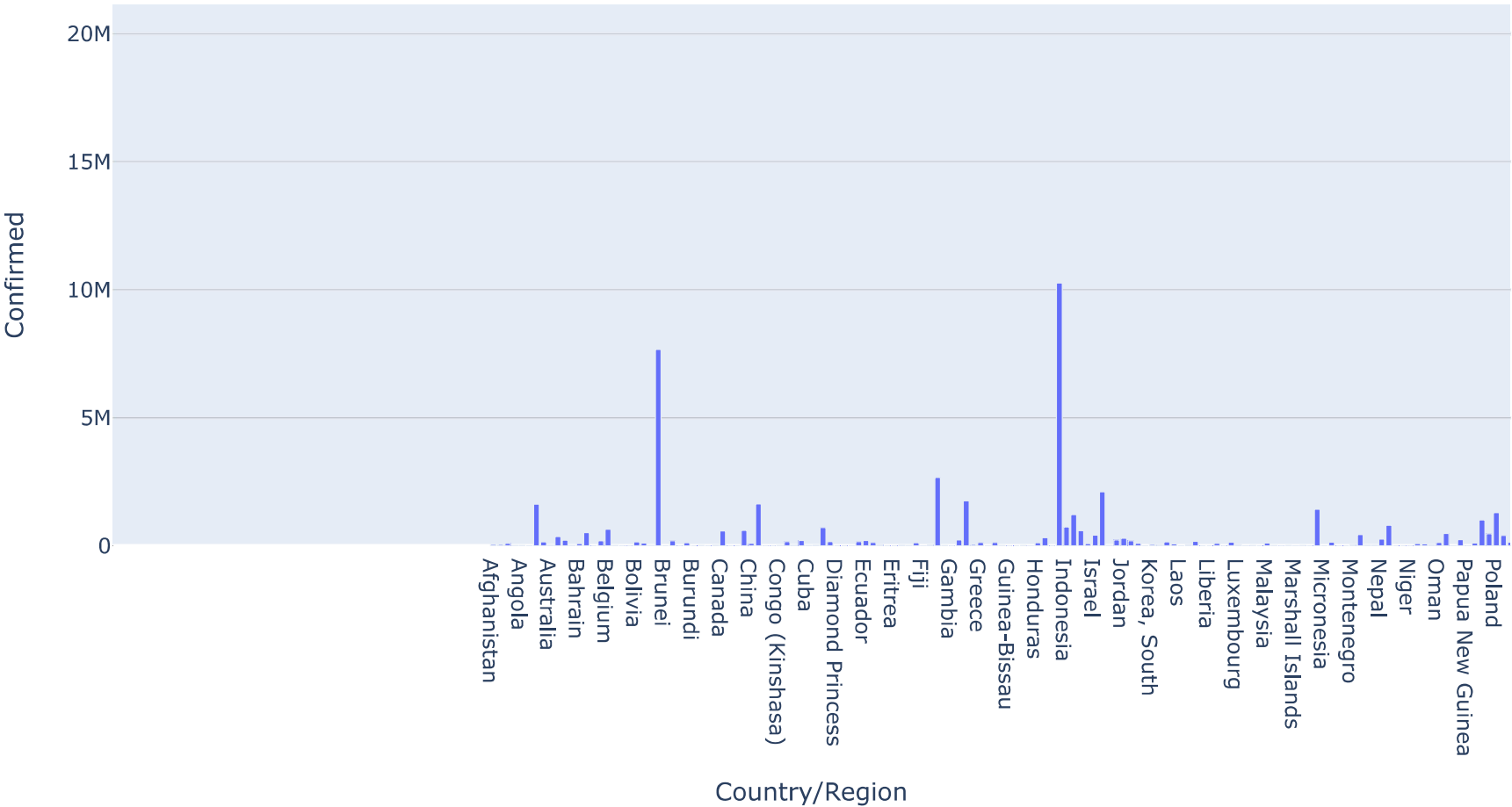
```
In [59]: fig=px.bar(total_confirmed_df,x='Country/Region',y='Confirmed')
fig.show()
```



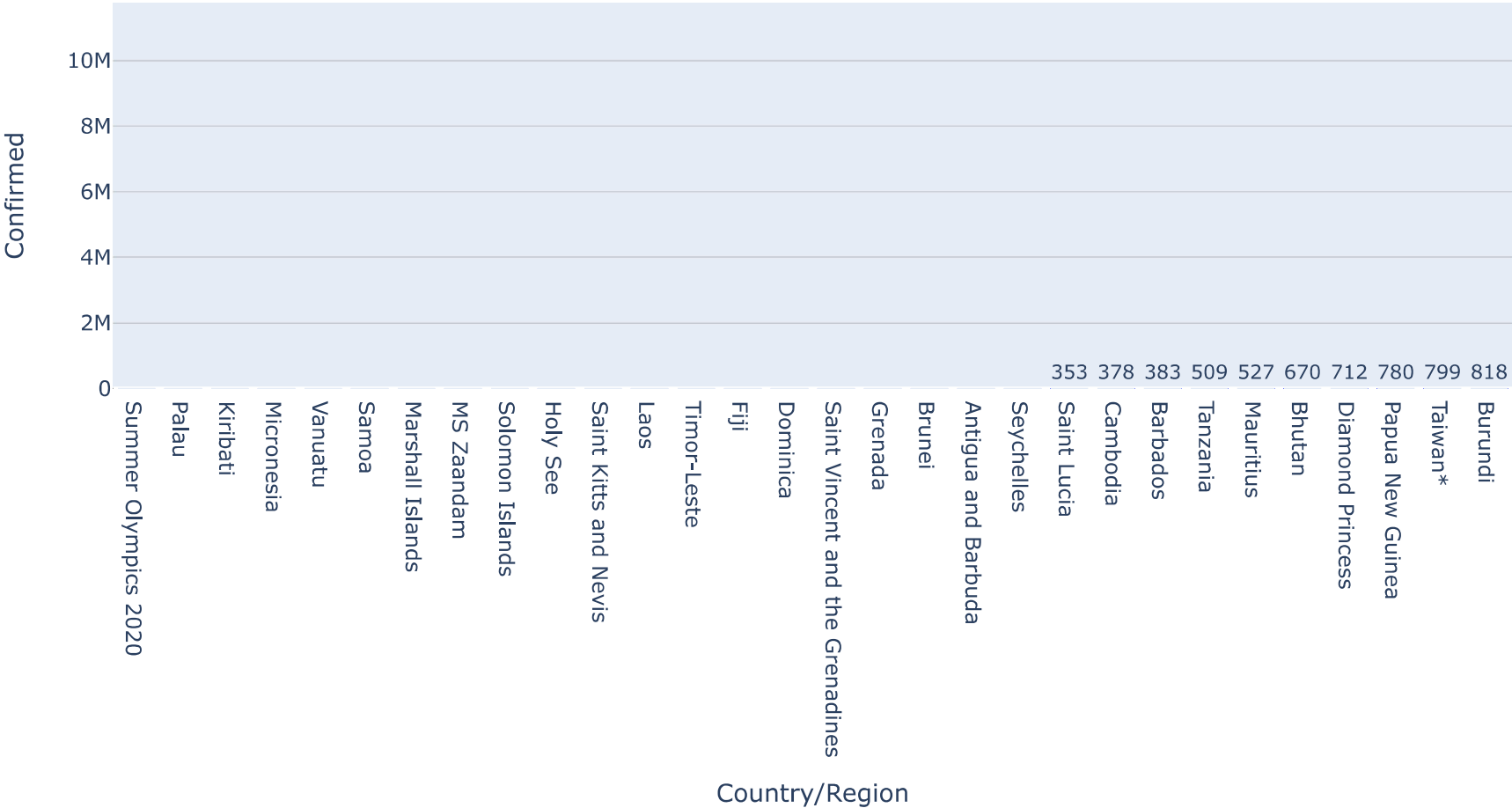
```
In [58]: fig.show()
```



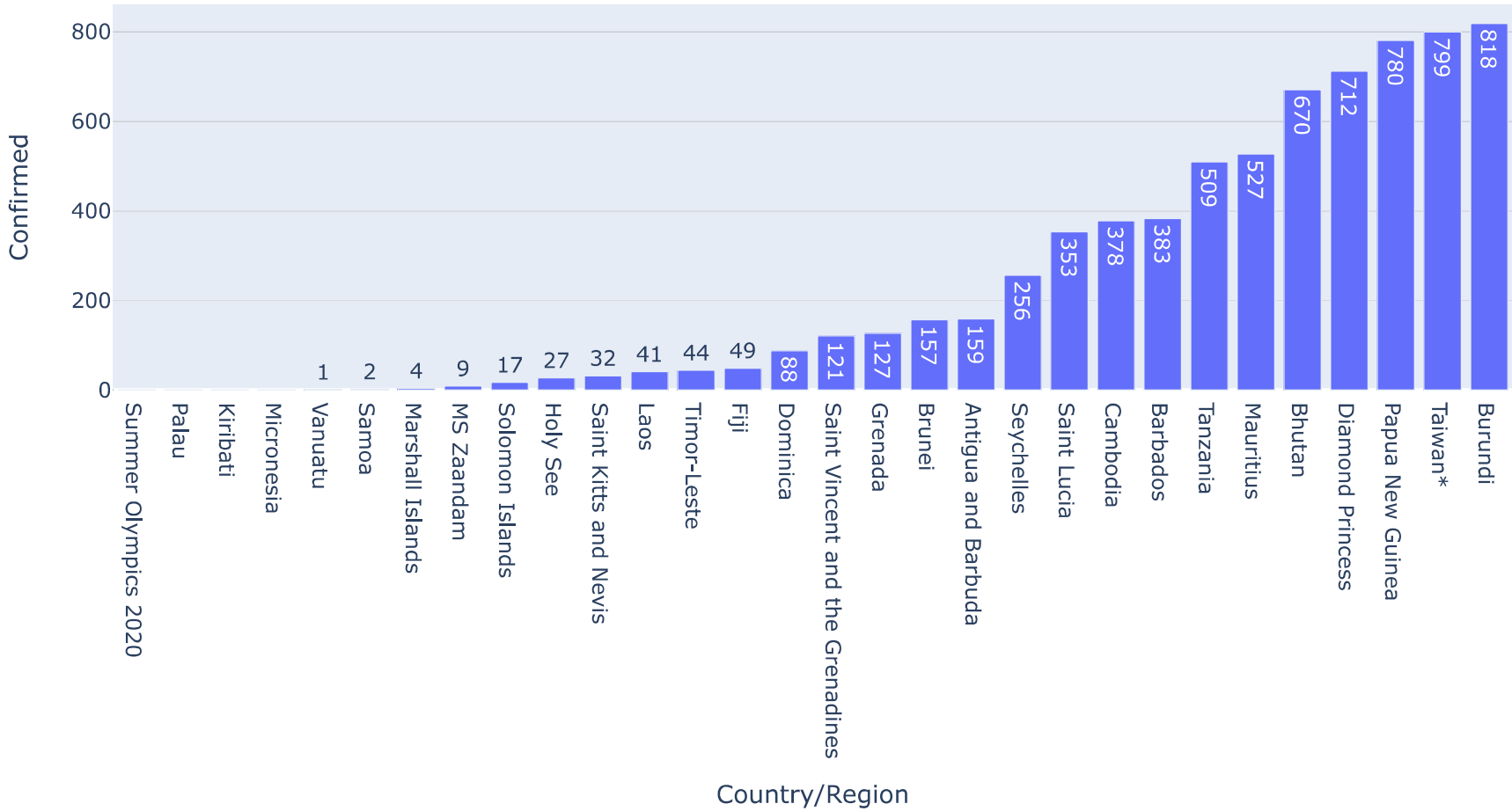
```
In [45]: fig.show()
```



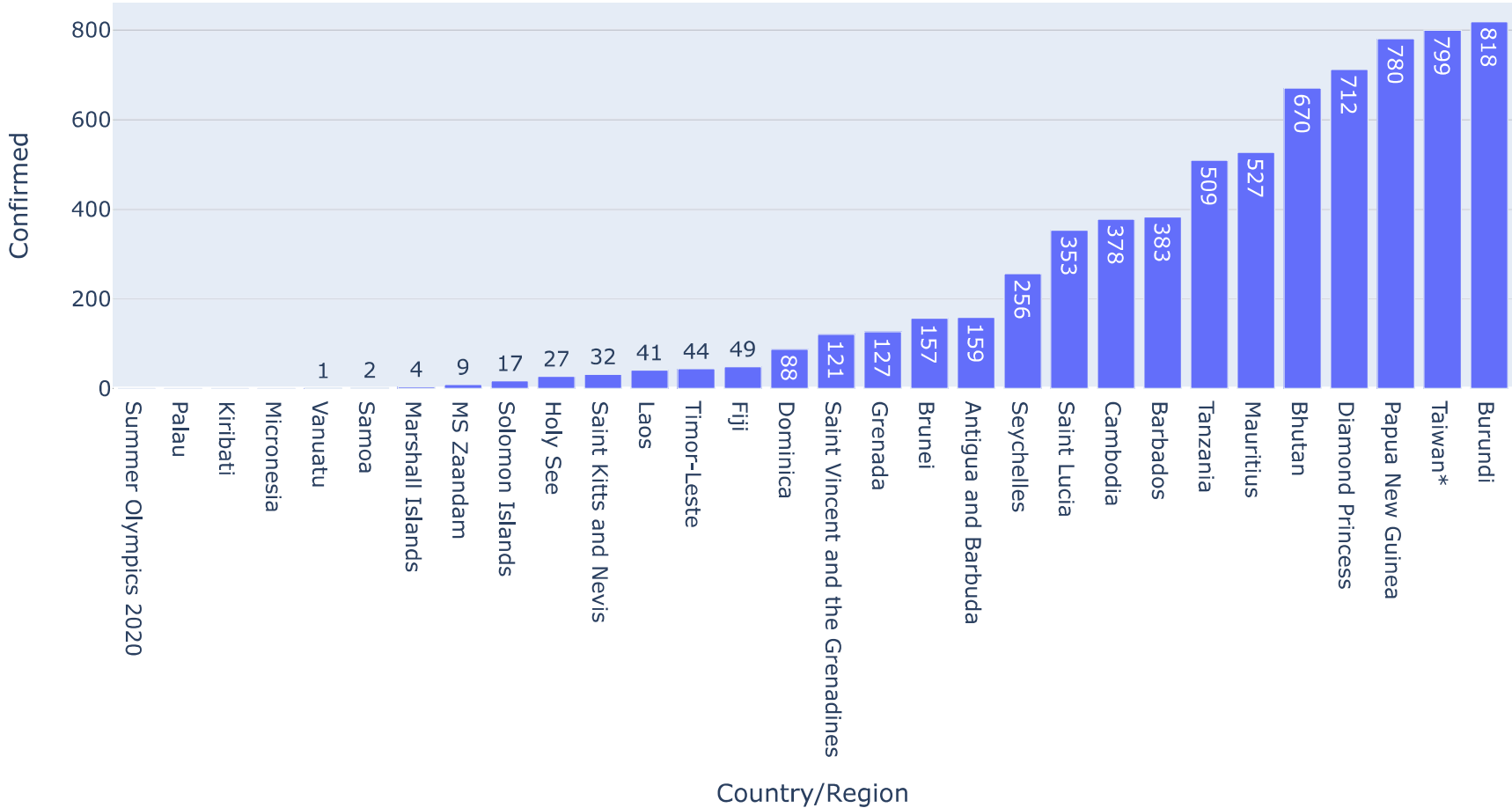
```
In [47]: fig=px.bar(total_confirmed_df.sort_values('Confirmed',ascending='False').head(30)
,x='Country/Region',y='Confirmed',text='Confirmed')
fig.show()
```



```
In [48]: fig=px.bar(total_confirmed_df.sort_values('Confirmed',ascending='False').head(30)
,x='Country/Region',y='Confirmed',text='Confirmed')
fig.show()
```




```
In [49]: fig=px.bar(total_confirmed_df.sort_values('Confirmed',ascending='True').head(30)
          ,x='Country/Region',y='Confirmed',text='Confirmed')
fig.show()
```

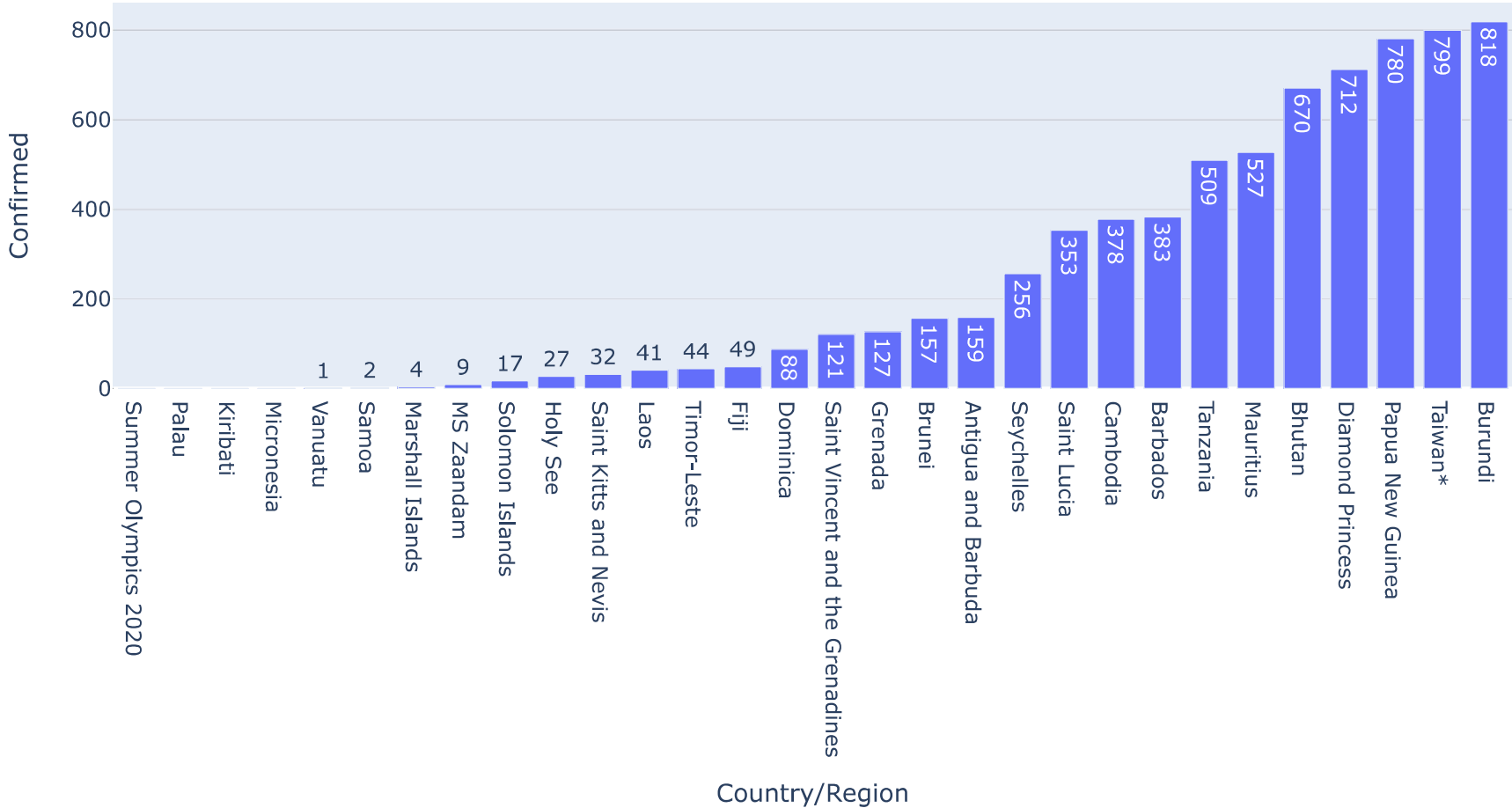


```
In [50]: ▶ fig=px.bar(total_confirmed_df.sort_values('Confirmed',descending='True').head(30)
          ,x='Country/Region',y='Confirmed',text='Confirmed')
fig.show()
```

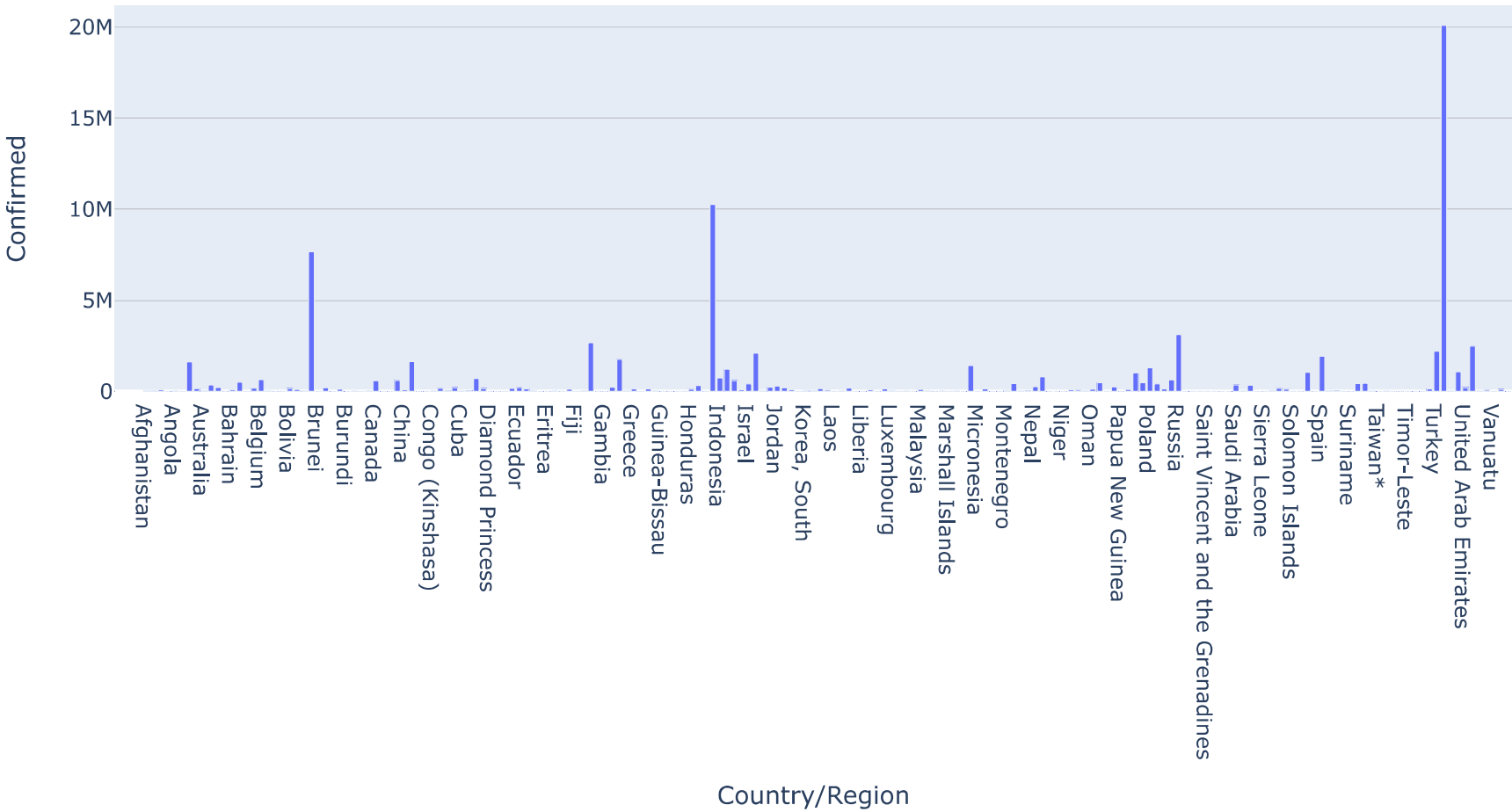
```
-----
TypeError                                Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_24968\162337786.py in <module>
----> 1 fig=px.bar(total_confirmed_df.sort_values('Confirmed',descending='True').head(30)
      2         ,x='Country/Region',y='Confirmed',text='Confirmed')
      3 fig.show()
```

TypeError: sort_values() got an unexpected keyword argument 'descending'

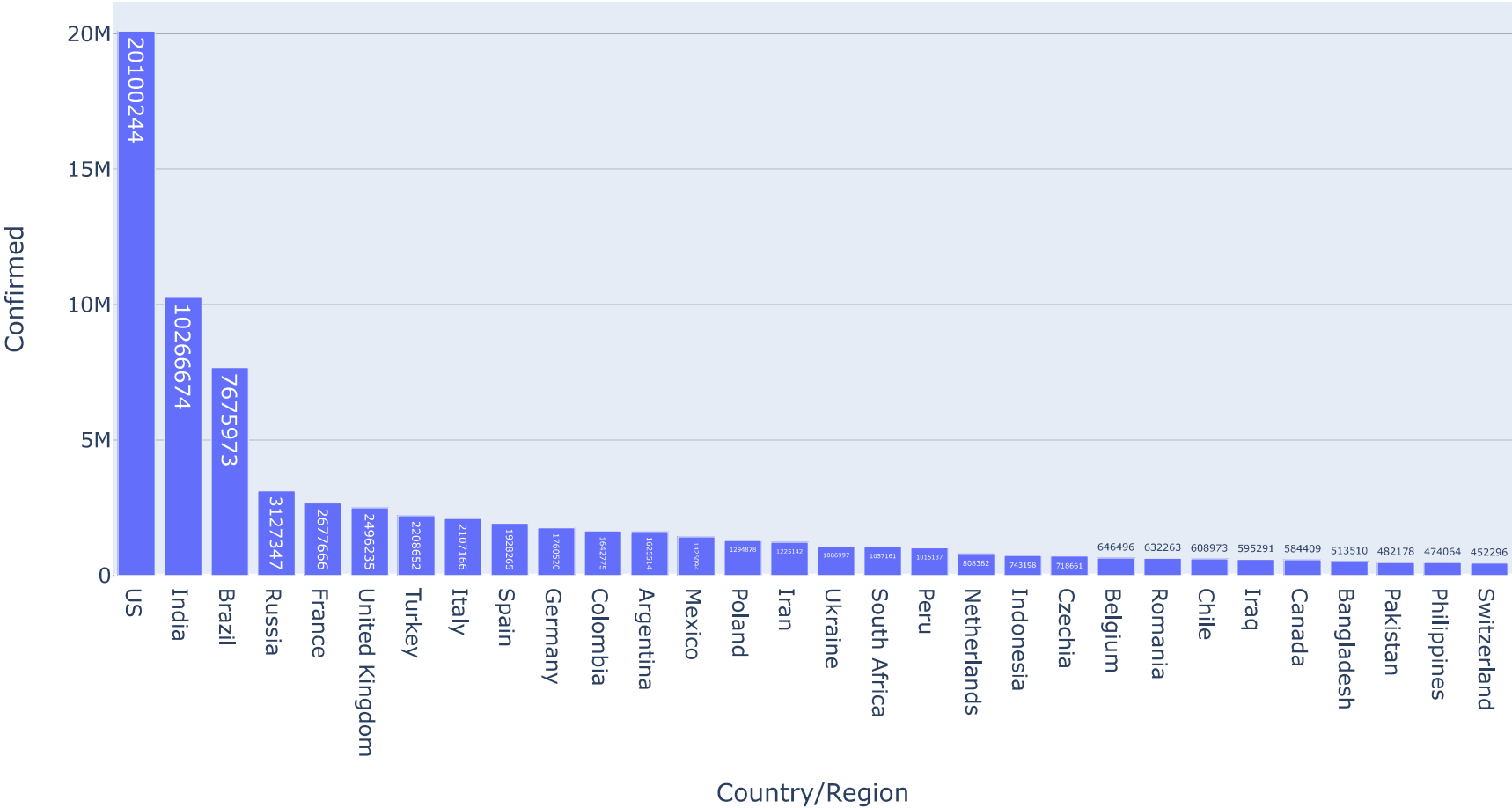
```
In [51]: fig=px.bar(total_confirmed_df.sort_values('Confirmed',ascending='False').head(30)
          ,x='Country/Region',y='Confirmed',text='Confirmed')
fig.show()
```



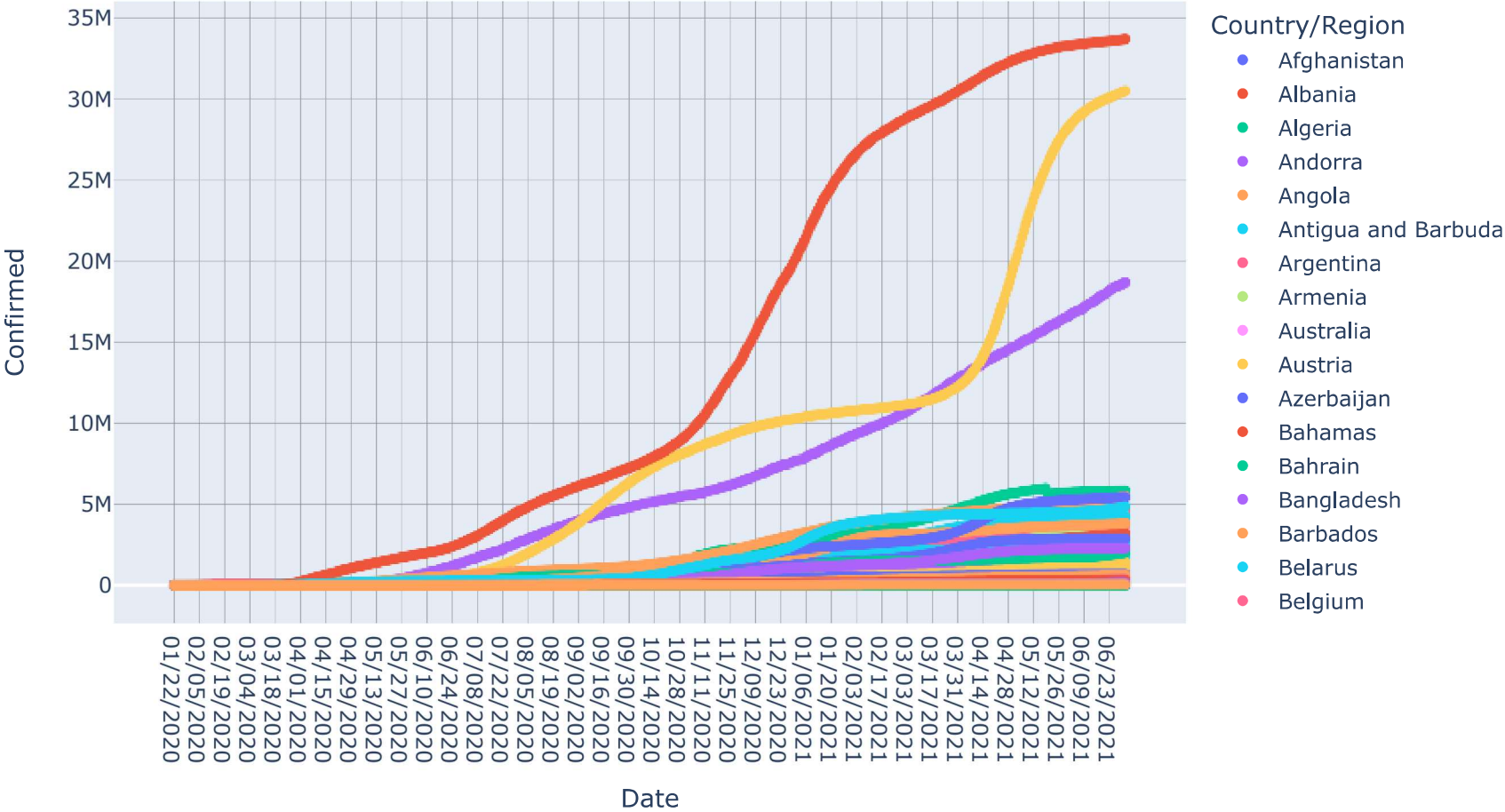
```
In [52]: fig=px.bar(total_confirmed_df,x='Country/Region',y='Confirmed')
fig.show()
```



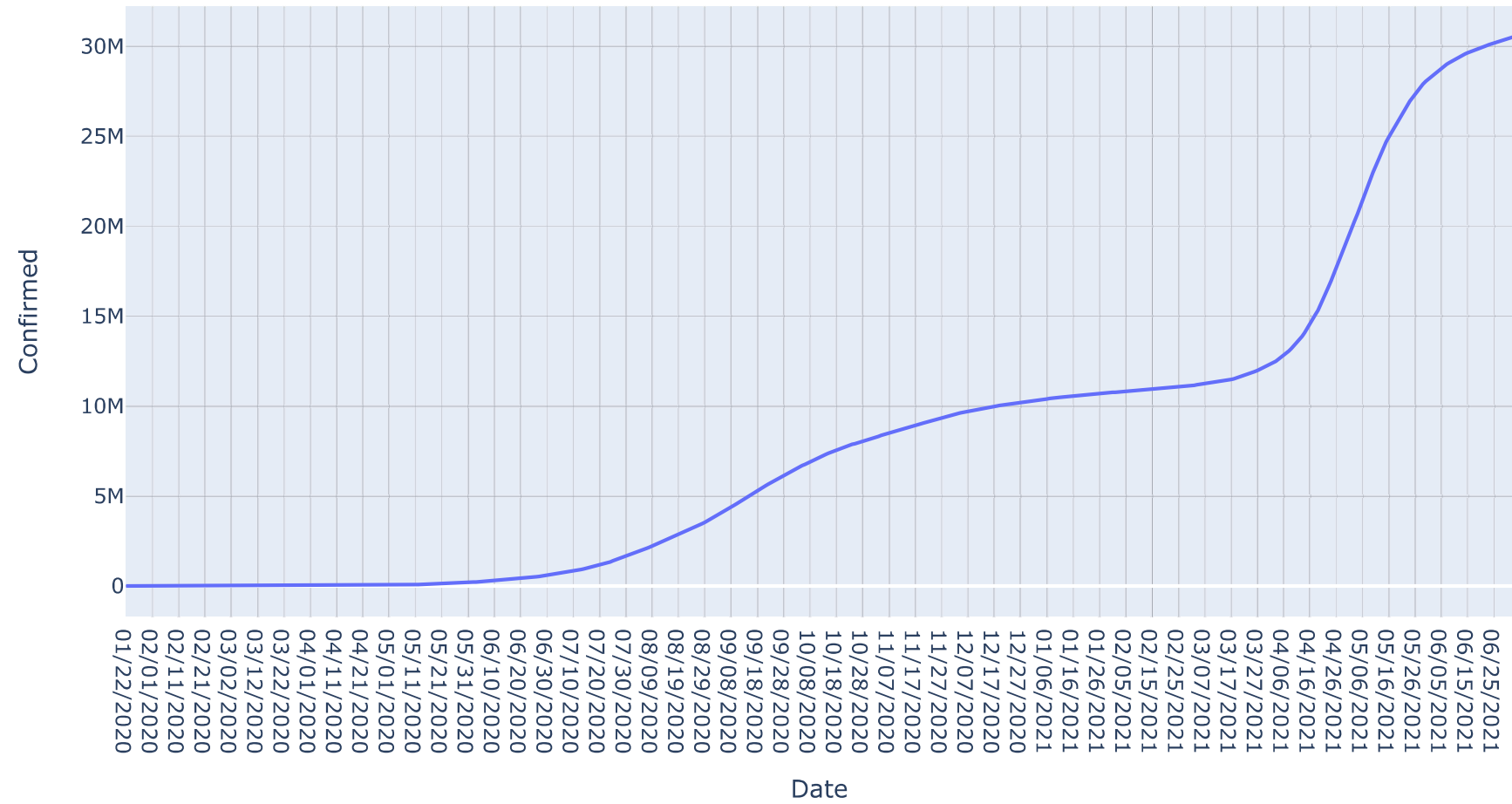
```
In [60]: fig=px.bar(total_confirmed_df.sort_values('Confirmed',ascending=False).head(30)
          ,x='Country/Region',y='Confirmed',text='Confirmed')
fig.show()
```



```
In [61]: fig2=px.scatter(confirmed_melt_df,x='Date',y='Confirmed',color='Country/Region')
fig2.show()
```



```
In [63]: fig3=px.line(confirmed_melt_df[confirmed_melt_df['Country/Region']=='India'],x='Date',y='Confirmed')
fig3.show()
```



```
In [ ]:
```

