In [1]:

```python
import pandas as pd


import matplotlib as mpl
import matplotlib.pyplot as plt

mpl.rcParams['figure.figsize'] = (20, 10)
mpl.rcParams['axes.grid'] = False
%matplotlib inline
```
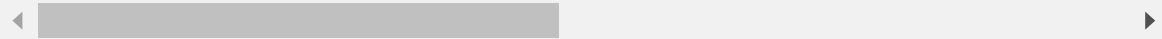
In [2]:

```python
covid_confirmed=pd.read_csv(r"C:\Users\hello\Desktop\python\covid\COVID-19-master\csse_
covid_19_data\csse_covid_19_time_series\time_series_19-covid-Confirmed.csv")
covid_confirmed
```

Out[2]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/
0	NaN	Thailand	15.0000	101.0000	2	3	5	7	
1	NaN	Japan	36.0000	138.0000	2	1	2	2	
2	NaN	Singapore	1.2833	103.8333	0	1	3	3	
3	NaN	Nepal	28.1667	84.2500	0	0	0	1	
4	NaN	Malaysia	2.5000	112.5000	0	0	0	3	
...	...	...	...	...	...	...	...	...	
482	NaN	Dominica	15.4150	-61.3710	0	0	0	0	
483	NaN	Grenada	12.1165	-61.6790	0	0	0	0	
484	NaN	Mozambique	-18.6657	35.5296	0	0	0	0	
485	NaN	Syria	34.8021	38.9968	0	0	0	0	
486	NaN	Timor-Leste	-8.8742	125.7275	0	0	0	0	

487 rows × 65 columns

In [3]:

```
covid_death=pd.read_csv(r"C:\Users\hello\Desktop\python\covid\COVID-19-master\csse_covi
d_19_data\csse_covid_19_time_series\time_series_19-covid-Deaths.csv")
covid_death
```

Out[3]:

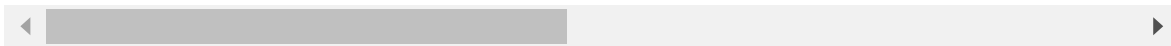|	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/
0	NaN	Thailand	15.0000	101.0000	0	0	0	0	
1	NaN	Japan	36.0000	138.0000	0	0	0	0	
2	NaN	Singapore	1.2833	103.8333	0	0	0	0	
3	NaN	Nepal	28.1667	84.2500	0	0	0	0	
4	NaN	Malaysia	2.5000	112.5000	0	0	0	0	
...	...	...	...	...	...	...	...	...	
482	NaN	Dominica	15.4150	-61.3710	0	0	0	0	
483	NaN	Grenada	12.1165	-61.6790	0	0	0	0	
484	NaN	Mozambique	-18.6657	35.5296	0	0	0	0	
485	NaN	Syria	34.8021	38.9968	0	0	0	0	
486	NaN	Timor-Leste	-8.8742	125.7275	0	0	0	0	

487 rows × 65 columns

◀ ◁▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ ▷

In [4]:

```
covid_recovered=pd.read_csv(r"C:\Users\hello\Desktop\python\covid\COVID-19-master\csse_
covid_19_data\csse_covid_19_time_series\time_series_19-covid-recovered.csv")
covid_recovered
```

Out[4]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/
0	NaN	Thailand	15.0000	101.0000	0	0	0	0	
1	NaN	Japan	36.0000	138.0000	0	0	0	0	
2	NaN	Singapore	1.2833	103.8333	0	0	0	0	
3	NaN	Nepal	28.1667	84.2500	0	0	0	0	
4	NaN	Malaysia	2.5000	112.5000	0	0	0	0	
...	...	...	...	...	...	...	...	...	
482	NaN	Dominica	15.4150	-61.3710	0	0	0	0	
483	NaN	Grenada	12.1165	-61.6790	0	0	0	0	
484	NaN	Mozambique	-18.6657	35.5296	0	0	0	0	
485	NaN	Syria	34.8021	38.9968	0	0	0	0	
486	NaN	Timor-Leste	-8.8742	125.7275	0	0	0	0	

487 rows × 65 columns

In [5]:

```
covid_recovered.head()
```

Out[5]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/
0	NaN	Thailand	15.0000	101.0000	0	0	0	0	
1	NaN	Japan	36.0000	138.0000	0	0	0	0	
2	NaN	Singapore	1.2833	103.8333	0	0	0	0	
3	NaN	Nepal	28.1667	84.2500	0	0	0	0	
4	NaN	Malaysia	2.5000	112.5000	0	0	0	0	

5 rows × 65 columns

In [6]:

```
covid_death.head()
```

Out[6]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/
0	NaN	Thailand	15.0000	101.0000	0	0	0	0	
1	NaN	Japan	36.0000	138.0000	0	0	0	0	
2	NaN	Singapore	1.2833	103.8333	0	0	0	0	
3	NaN	Nepal	28.1667	84.2500	0	0	0	0	
4	NaN	Malaysia	2.5000	112.5000	0	0	0	0	

5 rows × 65 columns

◀ ▶

In [7]:

```
covid_confirmed.head()
```

Out[7]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/
0	NaN	Thailand	15.0000	101.0000	2	3	5	7	
1	NaN	Japan	36.0000	138.0000	2	1	2	2	
2	NaN	Singapore	1.2833	103.8333	0	1	3	3	
3	NaN	Nepal	28.1667	84.2500	0	0	0	1	
4	NaN	Malaysia	2.5000	112.5000	0	0	0	3	

5 rows × 65 columns

◀ ▶

In [8]:

```
covid_confirmed=covid_confirmed.melt(id_vars=['Province/State','Country/Region','Lat',
'Long'])
covid_confirmed.head()
```

Out[8]:

	Province/State	Country/Region	Lat	Long	variable	value
0	NaN	Thailand	15.0000	101.0000	1/22/20	2
1	NaN	Japan	36.0000	138.0000	1/22/20	2
2	NaN	Singapore	1.2833	103.8333	1/22/20	0
3	NaN	Nepal	28.1667	84.2500	1/22/20	0
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0

In [9]:

```python
covid_death=covid_death.melt(id_vars=['Province/State','Country/Region','Lat','Long'])
covid_death.head()
```

Out[9]:

	Province/State	Country/Region	Lat	Long	variable	value
0	NaN	Thailand	15.0000	101.0000	1/22/20	0
1	NaN	Japan	36.0000	138.0000	1/22/20	0
2	NaN	Singapore	1.2833	103.8333	1/22/20	0
3	NaN	Nepal	28.1667	84.2500	1/22/20	0
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0

In [10]:

```python
covid_death=covid_death.rename({'variable':'Date','value': 'Death'}, axis='columns')
covid_death.head()
```

Out[10]:

	Province/State	Country/Region	Lat	Long	Date	Death
0	NaN	Thailand	15.0000	101.0000	1/22/20	0
1	NaN	Japan	36.0000	138.0000	1/22/20	0
2	NaN	Singapore	1.2833	103.8333	1/22/20	0
3	NaN	Nepal	28.1667	84.2500	1/22/20	0
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0

In [11]:

```python
covid_recovered=covid_recovered.melt(id_vars=['Province/State','Country/Region','Lat',
'Long']).rename({'variable':'Date','value': 'Recovered'}, axis='columns')
covid_recovered.head()
```

Out[11]:

	Province/State	Country/Region	Lat	Long	Date	Recovered
0	NaN	Thailand	15.0000	101.0000	1/22/20	0
1	NaN	Japan	36.0000	138.0000	1/22/20	0
2	NaN	Singapore	1.2833	103.8333	1/22/20	0
3	NaN	Nepal	28.1667	84.2500	1/22/20	0
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0

In [12]:

```
covid_confirmed=covid_confirmed.rename({'variable':'Date','value': 'confirmed'}, axis=
'columns')
covid_confirmed.head()
```

Out[12]:

	Province/State	Country/Region	Lat	Long	Date	confirmed
0	NaN	Thailand	15.0000	101.0000	1/22/20	2
1	NaN	Japan	36.0000	138.0000	1/22/20	2
2	NaN	Singapore	1.2833	103.8333	1/22/20	0
3	NaN	Nepal	28.1667	84.2500	1/22/20	0
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0

In [13]:

```
combined_covid = [covid_confirmed, covid_death, covid_recovered]
combined_covid
```

Out[13]:

```
[      Province/State Country/Region      Lat      Long      Date  confirme
d
 0              NaN        Thailand  15.0000  101.0000  1/22/20
2
 1              NaN           Japan  36.0000  138.0000  1/22/20
2
 2              NaN       Singapore   1.2833  103.8333  1/22/20
0
 3              NaN           Nepal  28.1667   84.2500  1/22/20
0
 4              NaN        Malaysia   2.5000  112.5000  1/22/20
0
 ...            ...             ...      ...       ...       ...
...
 29702          NaN        Dominica  15.4150  -61.3710  3/22/20
1
 29703          NaN         Grenada  12.1165  -61.6790  3/22/20
1
 29704          NaN      Mozambique -18.6657   35.5296  3/22/20
1
 29705          NaN           Syria  34.8021   38.9968  3/22/20
1
 29706          NaN     Timor-Leste  -8.8742  125.7275  3/22/20
1

 [29707 rows x 6 columns],
      Province/State Country/Region      Lat      Long      Date  Death
 0              NaN        Thailand  15.0000  101.0000  1/22/20      0
 1              NaN           Japan  36.0000  138.0000  1/22/20      0
 2              NaN       Singapore   1.2833  103.8333  1/22/20      0
 3              NaN           Nepal  28.1667   84.2500  1/22/20      0
 4              NaN        Malaysia   2.5000  112.5000  1/22/20      0
 ...            ...             ...      ...       ...       ...    ...
 29702          NaN        Dominica  15.4150  -61.3710  3/22/20      0
 29703          NaN         Grenada  12.1165  -61.6790  3/22/20      0
 29704          NaN      Mozambique -18.6657   35.5296  3/22/20      0
 29705          NaN           Syria  34.8021   38.9968  3/22/20      0
 29706          NaN     Timor-Leste  -8.8742  125.7275  3/22/20      0

 [29707 rows x 6 columns],
      Province/State Country/Region      Lat      Long      Date  Recovere
d
 0              NaN        Thailand  15.0000  101.0000  1/22/20
0
 1              NaN           Japan  36.0000  138.0000  1/22/20
0
 2              NaN       Singapore   1.2833  103.8333  1/22/20
0
 3              NaN           Nepal  28.1667   84.2500  1/22/20
0
 4              NaN        Malaysia   2.5000  112.5000  1/22/20
0
 ...            ...             ...      ...       ...       ...
...
 29702          NaN        Dominica  15.4150  -61.3710  3/22/20
0
 29703          NaN         Grenada  12.1165  -61.6790  3/22/20
0
 29704          NaN      Mozambique -18.6657   35.5296  3/22/20
```

```
0
 29705           NaN         Syria  34.8021   38.9968  3/22/20
0
 29706           NaN   Timor-Leste  -8.8742  125.7275  3/22/20
0

 [29707 rows x 6 columns]]
```

In [14]:

```python
combined_df = [covid_confirmed, covid_death, covid_recovered]
combined_df = [df.set_index(['Province/State','Country/Region','Lat','Long','Date']) fo
r df in combined_df]
combined_df=combined_df[0].join(combined_df[1:])
combined_df.head()
```

Out[14]:

Province/State	Country/Region	Lat	Long	Date	confirmed	Death	Recovered
NaN	Thailand	15.0000	101.0000	1/22/20	2	0	0
	Japan	36.0000	138.0000	1/22/20	2	0	0
	Singapore	1.2833	103.8333	1/22/20	0	0	0
	Nepal	28.1667	84.2500	1/22/20	0	0	0
	Malaysia	2.5000	112.5000	1/22/20	0	0	0

In [15]:

```python
combined_df=combined_df.reset_index()
combined_df.head()
```

Out[15]:

	Province/State	Country/Region	Lat	Long	Date	confirmed	Death	Recovered
0	NaN	Thailand	15.0000	101.0000	1/22/20	2	0	0
1	NaN	Japan	36.0000	138.0000	1/22/20	2	0	0
2	NaN	Singapore	1.2833	103.8333	1/22/20	0	0	0
3	NaN	Nepal	28.1667	84.2500	1/22/20	0	0	0
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0	0	0

In [16]:

```
combined_df[['Lat', 'Long','confirmed','Death','Recovered']] = combined_df[['Lat', 'Lon
g','confirmed','Death','Recovered']].apply(pd.to_numeric)
combined_df
```

Out[16]:

	Province/State	Country/Region	Lat	Long	Date	confirmed	Death	Recov
0	NaN	Thailand	15.0000	101.0000	1/22/20	2	0	
1	NaN	Japan	36.0000	138.0000	1/22/20	2	0	
2	NaN	Singapore	1.2833	103.8333	1/22/20	0	0	
3	NaN	Nepal	28.1667	84.2500	1/22/20	0	0	
4	NaN	Malaysia	2.5000	112.5000	1/22/20	0	0	
...	...	...	...	...	...	...	...	
29702	NaN	Dominica	15.4150	-61.3710	3/22/20	1	0	
29703	NaN	Grenada	12.1165	-61.6790	3/22/20	1	0	
29704	NaN	Mozambique	-18.6657	35.5296	3/22/20	1	0	
29705	NaN	Syria	34.8021	38.9968	3/22/20	1	0	
29706	NaN	Timor-Leste	-8.8742	125.7275	3/22/20	1	0	

29707 rows × 8 columns

In [17]:

```
combined_df[['Date']] = combined_df[['Date']].apply(pd.to_datetime)
```

In [18]:

```
combined_df.head()
```

Out[18]:

	Province/State	Country/Region	Lat	Long	Date	confirmed	Death	Recovered
0	NaN	Thailand	15.0000	101.0000	2020-01-22	2	0	0
1	NaN	Japan	36.0000	138.0000	2020-01-22	2	0	0
2	NaN	Singapore	1.2833	103.8333	2020-01-22	0	0	0
3	NaN	Nepal	28.1667	84.2500	2020-01-22	0	0	0
4	NaN	Malaysia	2.5000	112.5000	2020-01-22	0	0	0

In [19]:

```
combined_df.dtypes
```

Out[19]:

```
Province/State         object
Country/Region         object
Lat                   float64
Long                  float64
Date           datetime64[ns]
confirmed               int64
Death                   int64
Recovered               int64
dtype: object
```

In [20]:

```
combined_df=combined_df.rename({'Province/State':'State','Country/Region': 'Country'},
axis='columns')
```

In [21]:

```
combined_df.head()
```

Out[21]:

	State	Country	Lat	Long	Date	confirmed	Death	Recovered
0	NaN	Thailand	15.0000	101.0000	2020-01-22	2	0	0
1	NaN	Japan	36.0000	138.0000	2020-01-22	2	0	0
2	NaN	Singapore	1.2833	103.8333	2020-01-22	0	0	0
3	NaN	Nepal	28.1667	84.2500	2020-01-22	0	0	0
4	NaN	Malaysia	2.5000	112.5000	2020-01-22	0	0	0

In [22]:

```
combined_df.query("Country=='US' & State=='Washington'")
```

Out[22]:

	State	Country	Lat	Long	Date	confirmed	Death	Recovered
98	Washington	US	47.4009	-121.4905	2020-01-22	0	0	0
585	Washington	US	47.4009	-121.4905	2020-01-23	0	0	0
1072	Washington	US	47.4009	-121.4905	2020-01-24	0	0	0
1559	Washington	US	47.4009	-121.4905	2020-01-25	0	0	0
2046	Washington	US	47.4009	-121.4905	2020-01-26	0	0	0
...	...	...	...	...	...	...	...	...
27370	Washington	US	47.4009	-121.4905	2020-03-18	1014	55	0
27857	Washington	US	47.4009	-121.4905	2020-03-19	1376	74	0
28344	Washington	US	47.4009	-121.4905	2020-03-20	1524	83	0
28831	Washington	US	47.4009	-121.4905	2020-03-21	1793	94	0
29318	Washington	US	47.4009	-121.4905	2020-03-22	1996	95	0

61 rows × 8 columns

In [23]:

```
combined_df.query("Country=='China' & State=='Hubei'")
```

Out[23]:

	State	Country	Lat	Long	Date	confirmed	Death	Recovered
154	Hubei	China	30.9756	112.2707	2020-01-22	444	17	28
641	Hubei	China	30.9756	112.2707	2020-01-23	444	17	28
1128	Hubei	China	30.9756	112.2707	2020-01-24	549	24	31
1615	Hubei	China	30.9756	112.2707	2020-01-25	761	40	32
2102	Hubei	China	30.9756	112.2707	2020-01-26	1058	52	42
...	...	...	...	...	...	...	...	...
27426	Hubei	China	30.9756	112.2707	2020-03-18	67800	3122	56927
27913	Hubei	China	30.9756	112.2707	2020-03-19	67800	3130	57682
28400	Hubei	China	30.9756	112.2707	2020-03-20	67800	3133	58382
28887	Hubei	China	30.9756	112.2707	2020-03-21	67800	3139	58946
29374	Hubei	China	30.9756	112.2707	2020-03-22	67800	3144	59433

61 rows × 8 columns

In [24]:

```
combined_df['Active']=combined_df['confirmed'] - combined_df['Death'] - combined_df['Re
covered']
```

In [25]:

```python
combined_df.query("Country=='China' & State=='Hubei'")
```

Out[25]:

	State	Country	Lat	Long	Date	confirmed	Death	Recovered	Active
154	Hubei	China	30.9756	112.2707	2020-01-22	444	17	28	399
641	Hubei	China	30.9756	112.2707	2020-01-23	444	17	28	399
1128	Hubei	China	30.9756	112.2707	2020-01-24	549	24	31	494
1615	Hubei	China	30.9756	112.2707	2020-01-25	761	40	32	689
2102	Hubei	China	30.9756	112.2707	2020-01-26	1058	52	42	964
...	...	...	...	...	...	...	...	...	...
27426	Hubei	China	30.9756	112.2707	2020-03-18	67800	3122	56927	7751
27913	Hubei	China	30.9756	112.2707	2020-03-19	67800	3130	57682	6988
28400	Hubei	China	30.9756	112.2707	2020-03-20	67800	3133	58382	6285
28887	Hubei	China	30.9756	112.2707	2020-03-21	67800	3139	58946	5715
29374	Hubei	China	30.9756	112.2707	2020-03-22	67800	3144	59433	5223

61 rows × 9 columns

In [26]:

```python
max_df=combined_df[combined_df.Date==combined_df['Date'].max()]
```

In [27]:

```
max_df[['Country','Active']].sort_values('Active',ascending=False).head(10)
```

Out[27]:

	Country	Active
29236	Italy	46638
29231	Germany	24513
29238	Spain	24421
29319	US	15676
29377	France	13144
29375	Iran	12022
29251	Switzerland	7016
29376	Korea, South	5884
29623	United Kingdom	5337
29374	China	5223

In [28]:

```
max_df[['Country','confirmed']].groupby(['Country']).sum().sort_values('confirmed',asce
nding=False).head(10)
```

Out[28]:

|	confirmed
Country	
China	81397
Italy	59138
US	33272
Spain	28768
Germany	24873
Iran	21638
France	16176
Korea, South	8897
Switzerland	7245
United Kingdom	5741

In [29]:

```python
%matplotlib inline
mpl.rcParams['figure.figsize'] = (15, 10)
mpl.rcParams['axes.grid'] = False
combined_df[['Date','Country','State','Active']].query("Country=='China' & State=='Hubei'").plot(x='Date',y='Active',kind='bar', title="Corona Trend for Country China and State Hubei")
```

Out[29]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21e7fba2a90>
```
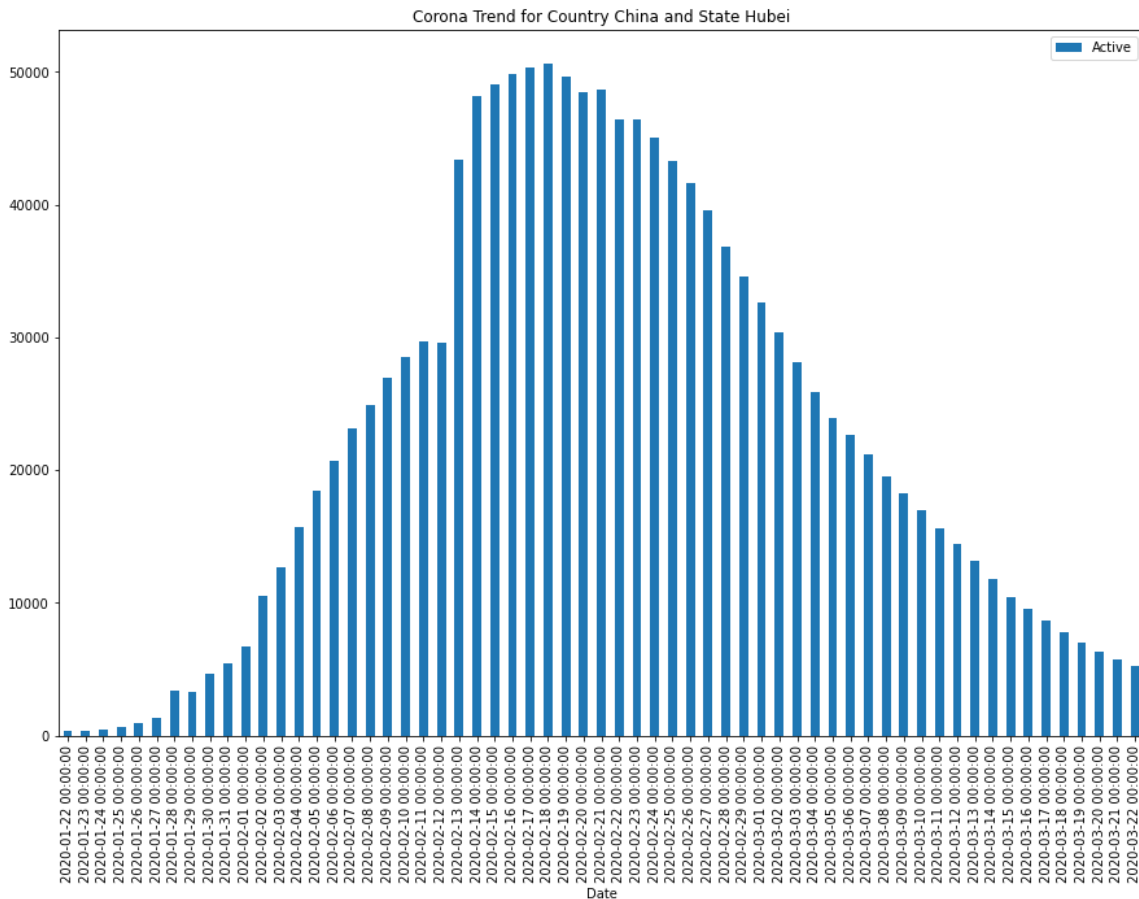
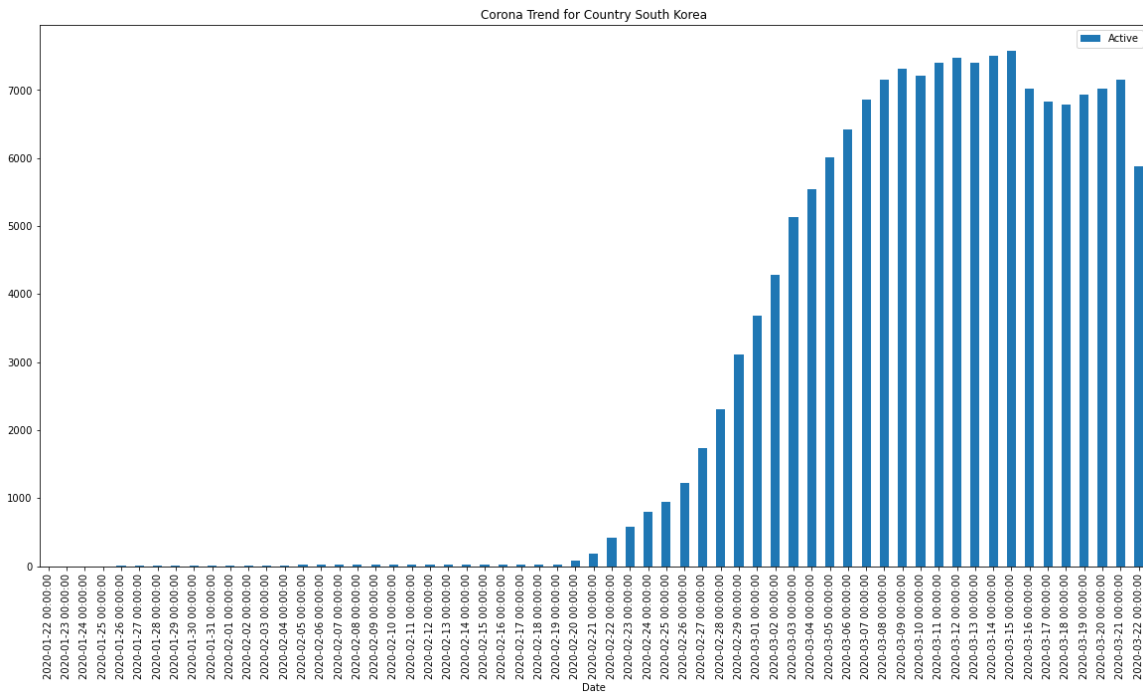In [30]:

```
%matplotlib inline
mpl.rcParams['figure.figsize'] = (20, 10)
mpl.rcParams['axes.grid'] = False
combined_df[['Date','Country','State','Active']].groupby(['Date','Country']).sum().query("Country=='China'").reset_index().plot(x='Date',y='Active',kind='bar',title="Corona Trend for Country China")
```

Out[30]:
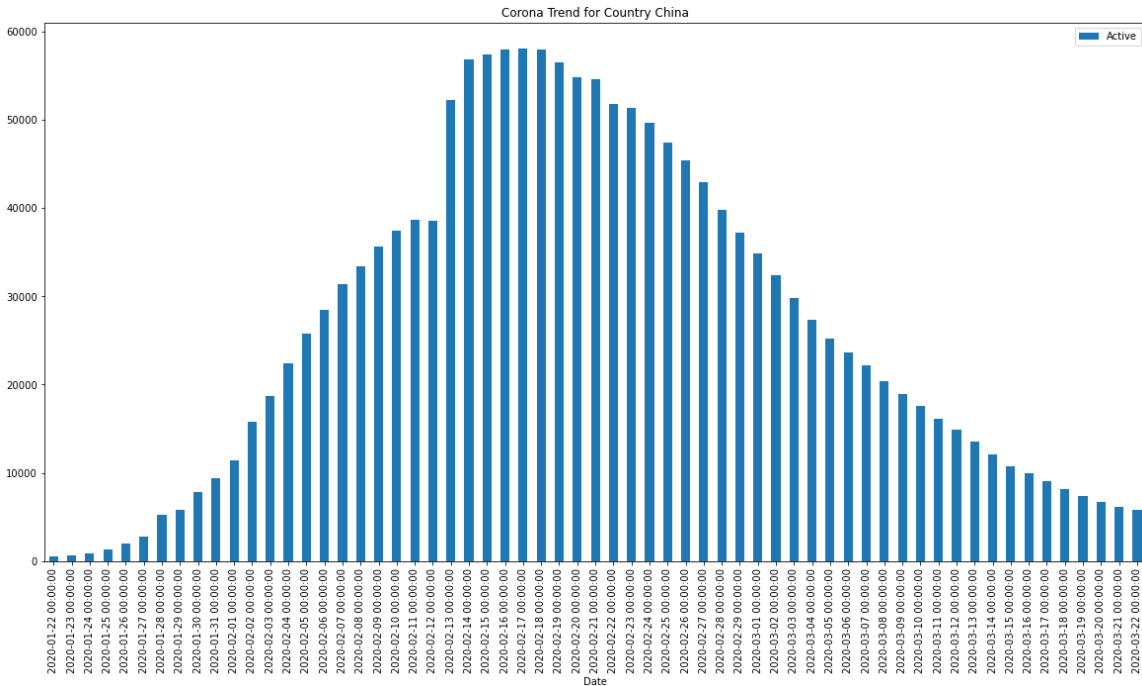
```
<matplotlib.axes._subplots.AxesSubplot at 0x21e7fd73a90>
```

In [31]:

```python
%matplotlib inline
mpl.rcParams['figure.figsize'] = (20, 10)
mpl.rcParams['axes.grid'] = False
combined_df[['Date','Country','State','Active']].query("Country=='Korea, South'").plot(
x='Date',y='Active',kind='box',title="Corona Trend for Country South Korea")
```
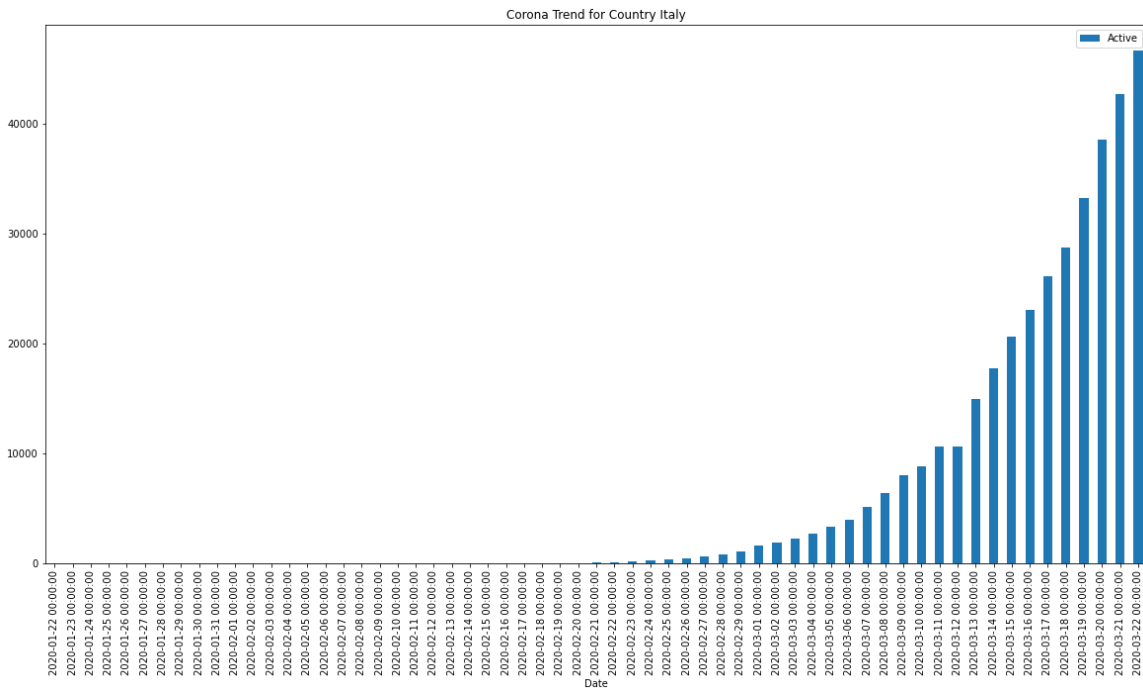
Out[31]:
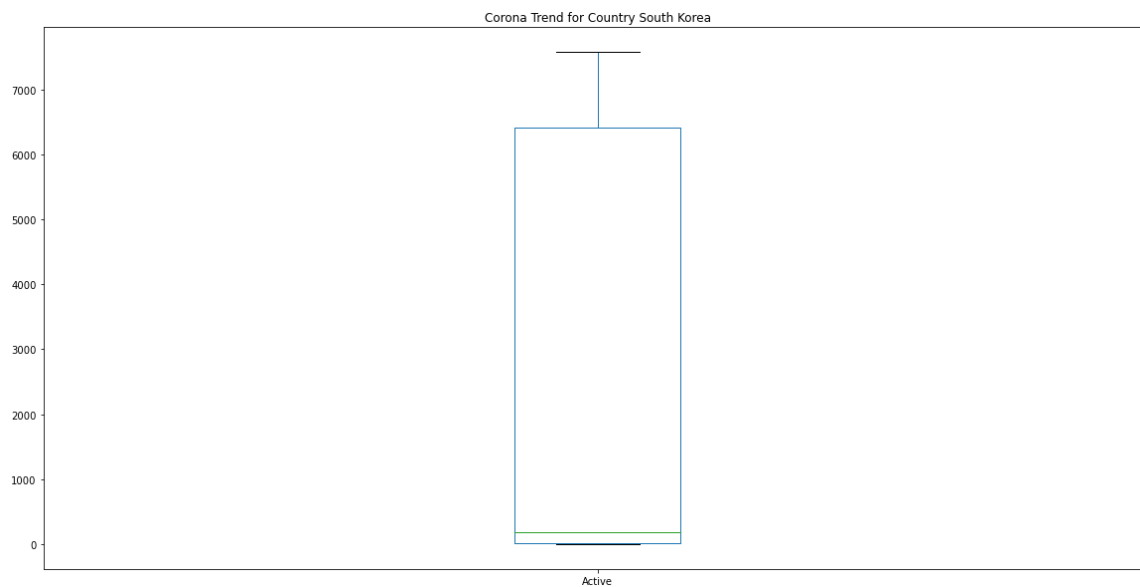
```
<matplotlib.axes._subplots.AxesSubplot at 0x21e7fcde6d0>
```

In [32]:

```python
%matplotlib inline
mpl.rcParams['figure.figsize'] = (20, 10)
mpl.rcParams['axes.grid'] = False
combined_df[['Date','Country','State','Active']].query("Country=='Korea, South'").plot(
x='Date',y='Active',kind='bar',title="Corona Trend for Country South Korea")
```

Out[32]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21e021d2370>
```

In [33]:

```python
%matplotlib inline
mpl.rcParams['figure.figsize'] = (20, 10)
mpl.rcParams['axes.grid'] = False
combined_df[['Date','Country','State','Active']].query("Country=='Italy'").plot(x='Date',y='Active',kind='bar', title="Corona Trend for Country Italy")
```
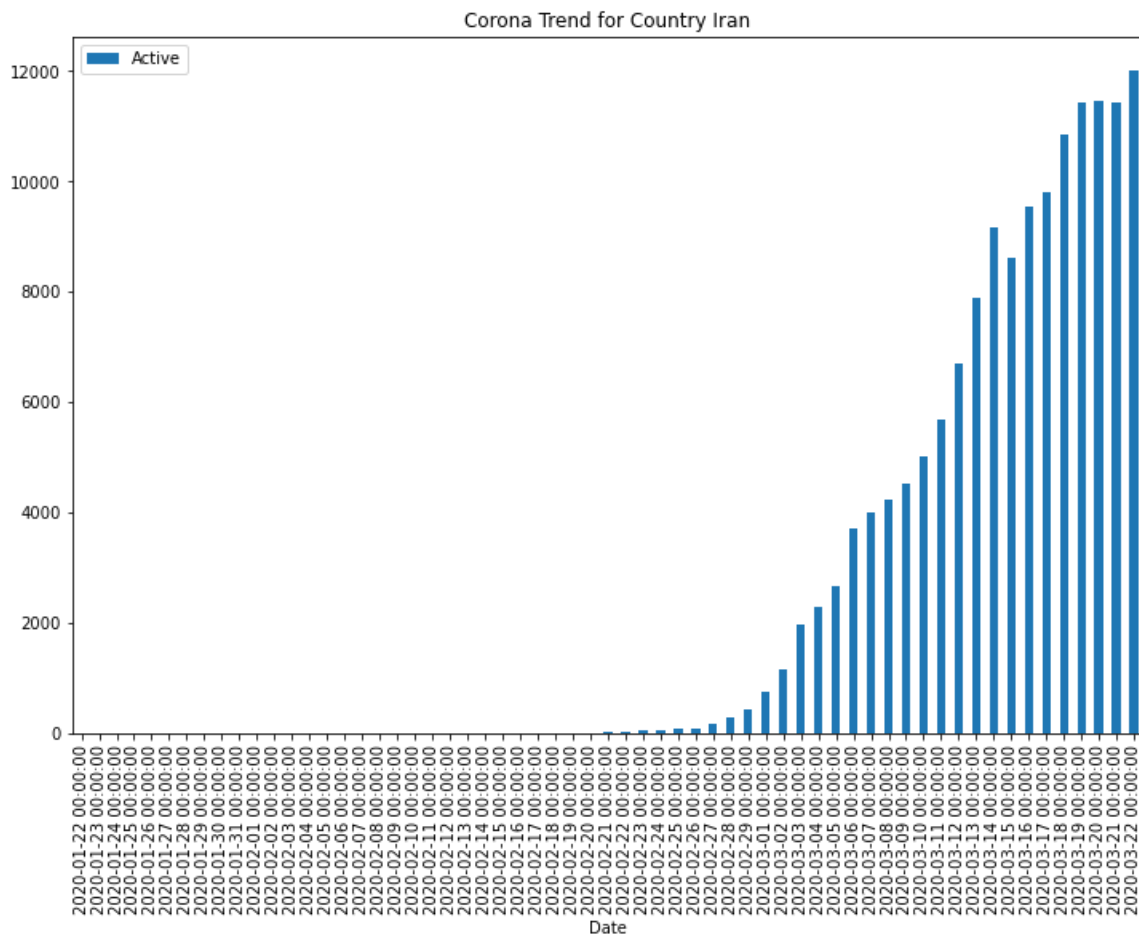
Out[33]:

<matplotlib.axes._subplots.AxesSubplot at 0x21e0233aa30>

Corona Trend for Country Italy

In [34]:

```python
%matplotlib inline
mpl.rcParams['figure.figsize'] = (12, 8)
mpl.rcParams['axes.grid'] = False
combined_df[['Date','Country','State','Active']].query("Country=='Iran'").plot(x='Date'
,y='Active',kind='bar', title="Corona Trend for Country Iran")
```

Out[34]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21e026ff5e0>
```


Corona Trend for Country Iran

In [ ]: