

# Covid-19 dashboard final

May 30, 2020

```
[1]: # importing libraries
import pandas as pd
import numpy as np
```

```
[2]: # loading data from the source:
confirmed_df = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/
    COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/
    time_series_covid19_confirmed_global.csv")
death_df = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/
    COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/
    time_series_covid19_deaths_global.csv")
recovered_df = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/
    COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/
    time_series_covid19_recovered_global.csv")
country_df = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/
    COVID-19/web-data/data/cases_country.csv")
```

```
[3]: confirmed_df.head()
```

```
[3]: Province/State Country/Region    Lat    Long  1/22/20  1/23/20  1/24/20  \
0      NaN      Afghanistan  33.0000  65.0000      0      0      0
1      NaN      Albania    41.1533  20.1683      0      0      0
2      NaN      Algeria    28.0339   1.6596      0      0      0
3      NaN      Andorra    42.5063   1.5218      0      0      0
4      NaN      Angola    -11.2027  17.8739      0      0      0

    1/25/20  1/26/20  1/27/20  ...  5/20/20  5/21/20  5/22/20  5/23/20  \
0          0          0          0  ...    8145    8676    9216    9998
1          0          0          0  ...     964     969     981     989
2          0          0          0  ...    7542    7728    7918    8113
3          0          0          0  ...     762     762     762     762
4          0          0          0  ...      52      58      60      61

    5/24/20  5/25/20  5/26/20  5/27/20  5/28/20  5/29/20
0    10582    11173    11831    12456    13036    13659
1       998     1004     1029     1050     1076     1099
2     8306     8503     8697     8857     8997     9134
```

3	762	763	763	763	763	764
4	69	70	70	71	74	81

[5 rows x 133 columns]

```
[4]: #death_df.head()
```

```
[5]: #recovered_df.head()
```

```
[6]: #country_df.head()
```

```
[7]: # data cleaning
```

```
# converting df column names to lowercase
```

```
country_df.columns = map(str.lower, country_df.columns)
confirmed_df.columns = map(str.lower, confirmed_df.columns)
death_df.columns = map(str.lower, death_df.columns)
recovered_df.columns = map(str.lower, recovered_df.columns)
```

```
[8]: # changing province/state to state and country/region to country
```

```
confirmed_df = confirmed_df.rename(columns={'province/state': 'state', 'country/
↪region': 'country'})
recovered_df = confirmed_df.rename(columns={'province/state': 'state', 'country/
↪region': 'country'})
death_df = death_df.rename(columns={'province/state': 'state', 'country/region':
↪ 'country'})
country_df = country_df.rename(columns={'country_region': 'country'})
```

```
[9]: #Sorting confirmed cases using descending order
```

```
sorted_country_df = country_df.sort_values('confirmed', ascending= False)
```

```
[11]: ## conda install -c https://conda.anaconda.org/plotly plotly (use this if
↪plotply is not installed in your system)
```

```
import plotly.express as px
```

```
[14]: fig= px.bar(
    sorted_country_df.head(10),
    x = "country",
    y = "confirmed",
    title= "Top 10 worst affected countries", # the axis names
    color = "country"
)
fig.update_layout(template='plotly_dark')
fig.show()
```

```
[15]: country_df.head()
```

```
[15]:
```

	country	last_update	lat	long_	confirmed	deaths	\
0	Australia	2020-05-30 08:32:43	-25.0000	133.0000	7185.0	103.0	
1	Austria	2020-05-30 08:32:43	47.5162	14.5501	16655.0	668.0	
2	Canada	2020-05-30 08:32:43	60.0010	-95.0010	90909.0	7063.0	
3	China	2020-05-30 08:32:43	30.5928	114.3055	84123.0	4638.0	
4	Denmark	2020-05-30 08:32:43	56.2639	9.5018	11793.0	568.0	

  

	recovered	active	incident_rate	people_tested	people_hospitalized	\
0	6606.0	476.0	28.221071	NaN	NaN	
1	15347.0	640.0	184.924054	NaN	NaN	
2	47905.0	35941.0	240.146121	NaN	NaN	
3	79383.0	102.0	5.988782	NaN	NaN	
4	10438.0	787.0	203.601290	NaN	NaN	

  

	mortality_rate	uid	iso3
0	1.433542	36	AUS
1	4.010808	40	AUT
2	7.769308	124	CAN
3	5.513355	156	CHN
4	4.816417	208	DNK

```
[16]: country_df.drop(['lat', 'long_', 'last_update', 'incident_rate',  
→ 'people_tested', 'people_hospitalized', 'mortality_rate', 'uid', 'iso3'],  
→ axis=1).sort_values('confirmed', ascending = False).reset_index(drop=True).  
→ style.bar(align='left', width=98, color='green')
```

```
[16]: <pandas.io.formats.style.Styler at 0x7f148a284a90>
```

```
[ ]:
```