In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
confirmed_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_19-covid-Confirmed.csv')
```

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

Out[3]:

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20
0	NaN	Thailand	15.0000	101.0000	2	3	5	7	8
1	NaN	Japan	36.0000	138.0000	2	1	2	2	4
2	NaN	Singapore	1.2833	103.8333	0	1	3	3	4
3	NaN	Nepal	28.1667	84.2500	0	0	0	1	1
4	NaN	Malaysia	2.5000	112.5000	0	0	0	3	4

5 rows × 66 columns

In [4]:
```
cols = confirmed_df.keys()
```

In [5]:
```
cols
```

Out[5]:
```
Index(['Province/State', 'Country/Region', 'Lat', 'Long', '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', '1/31/20', '2/1/20', '2/2/20', '2/3/20', '2/4/20', '2/5/20',
       '2/6/20', '2/7/20', '2/8/20', '2/9/20', '2/10/20', '2/11/20', '2/12/20',
       '2/13/20', '2/14/20', '2/15/20', '2/16/20', '2/17/20', '2/18/20',
       '2/19/20', '2/20/20', '2/21/20', '2/22/20', '2/23/20', '2/24/20',
       '2/25/20', '2/26/20', '2/27/20', '2/28/20', '2/29/20', '3/1/20',
       '3/2/20', '3/3/20', '3/4/20', '3/5/20', '3/6/20', '3/7/20', '3/8/20',
       '3/9/20', '3/10/20', '3/11/20', '3/12/20', '3/13/20', '3/14/20',
       '3/15/20', '3/16/20', '3/17/20', '3/18/20', '3/19/20', '3/20/20',
       '3/21/20', '3/22/20', '3/23/20'],
      dtype='object')
```

In [6]:
```
confirmed_dates = confirmed_df.loc[:, cols[4]:cols[-1]]
```

In [7]:
```
confirmed_dates
```

Out[7]:

	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	1/31/20	...
0	2	3	5	7	8	8	14	14	14	19	...
1	2	1	2	2	4	4	7	7	11	15	...
2	0	1	3	3	4	5	7	7	10	13	...
3	0	0	0	1	1	1	1	1	1	1	...
4	0	0	0	3	4	4	4	7	8	8	...
...	...	...	...	...	...	...	...	...	...	...	...
496	0	0	0	0	0	0	0	0	0	0	...
497	0	0	0	0	0	0	0	0	0	0	...
498	0	0	0	0	0	0	0	0	0	0	...
499	0	0	0	0	0	0	0	0	0	0	...
500	0	0	0	0	0	0	0	0	0	0	...

501 rows × 62 columns

In [8]:
```
dates = confirmed_dates.keys()
```

In [9]:
```
deaths_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_19-covid-Deaths.csv')
```

In [10]:
```
deaths_numbers = deaths_df.loc[:, cols[4]:cols[-1]]
```

In [11]:
```
deaths_df
```

Out[11]:

	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	
...	...	...	...	...	...	...	...	...	
496	NaN	Jersey	49.1900	-2.1100	0	0	0	0	
497	NaN	Puerto Rico	18.2000	-66.5000	0	0	0	0	
498	NaN	Republic of the Congo	-1.4400	15.5560	0	0	0	0	
499	NaN	The Bahamas	24.2500	-76.0000	0	0	0	0	
500	NaN	The Gambia	13.4667	-16.6000	0	0	0	0	

501 rows × 66 columns

In [12]:
```
recoveries_df = pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_19-covid-Recovered.csv')
```

In [13]:
```
recoveries_numbers = recoveries_df.loc[:, cols[4]:cols[-1]]
```

In [14]:
```
recoveries_df
```

Out[14]:

	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	
...	...	...	...	...	...	...	...	...	
496	NaN	Jersey	49.1900	-2.1100	0	0	0	0	
497	NaN	Puerto Rico	18.2000	-66.5000	0	0	0	0	
498	NaN	Republic of the Congo	-1.4400	15.5560	0	0	0	0	
499	NaN	The Bahamas	24.2500	-76.0000	0	0	0	0	
500	NaN	The Gambia	13.4667	-16.6000	0	0	0	0	

501 rows × 66 columns

In [ ]:
```
```

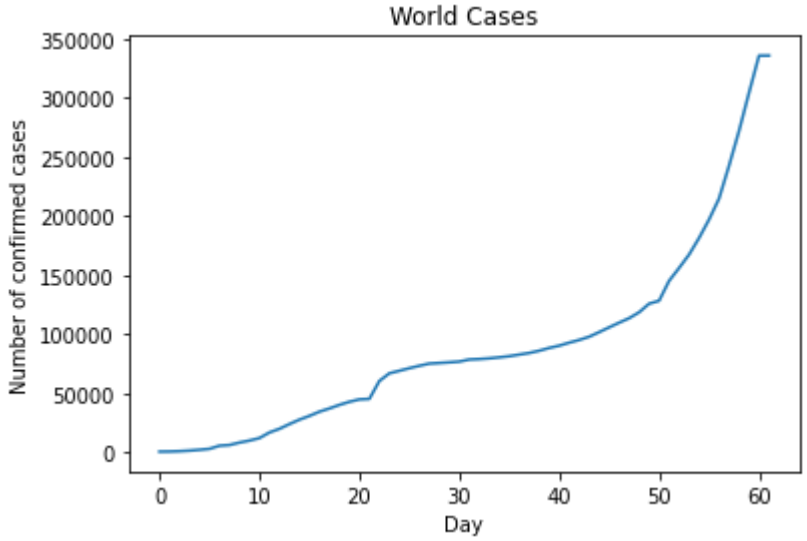
World Cases

In [15]:
```
world_cases = []
total_deaths = []
mortality_rate = []
total_recovered = []
for i in dates:
    confirmed_sum = confirmed_dates[i].sum()
    death_sum = deaths_numbers[i].sum()
    recovered_sum = recoveries_numbers[i].sum()
    world_cases.append(confirmed_sum)
    total_deaths.append(death_sum)
    mortality_rate.append(death_sum/confirmed_sum)
    total_recovered.append(recovered_sum)
```

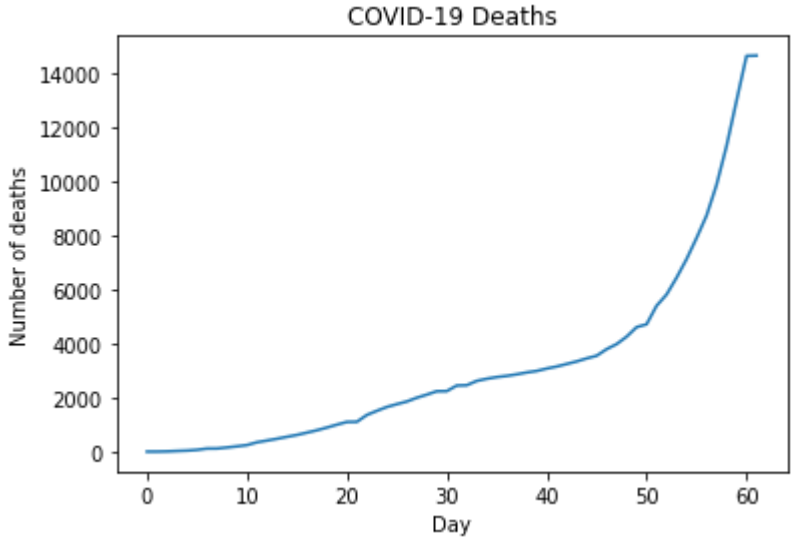
In [16]:
```
days_since_1_22 = np.array([i for i in range(len(dates))]).reshape(-1
, 1)
world_cases = np.array(world_cases).reshape(-1, 1)
total_deaths = np.array(total_deaths).reshape(-1, 1)
total_recovered = np.array(total_recovered).reshape(-1, 1)
```

In [17]:
```
import matplotlib.pyplot as plt
%matplotlib inline
```

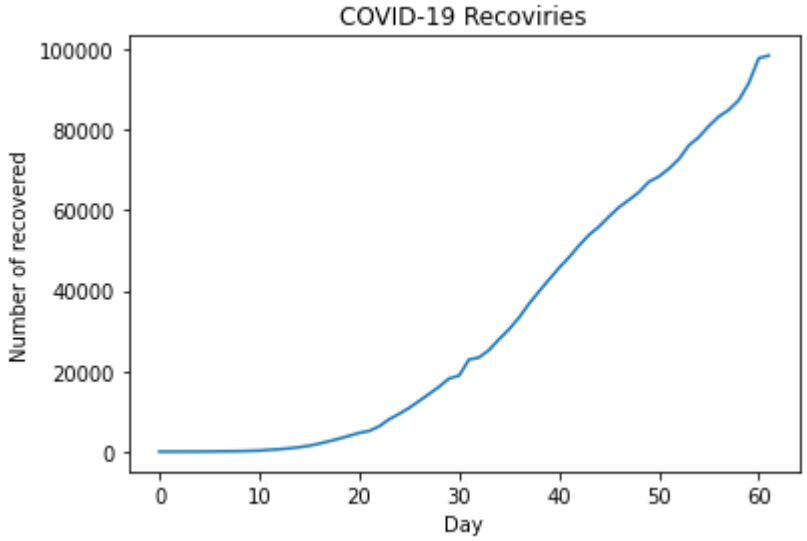
In [18]:
```
plt.plot(world_cases)
plt.title("World Cases")
plt.xlabel("Day")
plt.ylabel("Number of confirmed cases")
plt.show()
```



In [19]:
```
plt.plot(total_deaths)
plt.title("COVID-19 Deaths")
plt.xlabel("Day")
plt.ylabel("Number of deaths")
plt.show()
```



In [20]:
```
plt.plot(total_recovered)
plt.title("COVID-19 Recoviries")
plt.xlabel("Day")
plt.ylabel("Number of recovered")
plt.show()
```



In [ ]:

In [21]:
```
confirmed_df.keys()
```

Out[21]:
```
Index(['Province/State', 'Country/Region', 'Lat', 'Long', '1/22/20',
'1/23/20',
       '1/24/20', '1/25/20', '1/26/20', '1/27/20', '1/28/20', '1/29/2
0',
       '1/30/20', '1/31/20', '2/1/20', '2/2/20', '2/3/20', '2/4/20',
'2/5/20',
       '2/6/20', '2/7/20', '2/8/20', '2/9/20', '2/10/20', '2/11/20',
'2/12/20',
       '2/13/20', '2/14/20', '2/15/20', '2/16/20', '2/17/20', '2/18/2
0',
       '2/19/20', '2/20/20', '2/21/20', '2/22/20', '2/23/20', '2/24/2
0',
       '2/25/20', '2/26/20', '2/27/20', '2/28/20', '2/29/20', '3/1/2
0',
       '3/2/20', '3/3/20', '3/4/20', '3/5/20', '3/6/20', '3/7/20',
'3/8/20',
       '3/9/20', '3/10/20', '3/11/20', '3/12/20', '3/13/20', '3/14/2
0',
       '3/15/20', '3/16/20', '3/17/20', '3/18/20', '3/19/20', '3/20/2
0',
       '3/21/20', '3/22/20', '3/23/20'],
      dtype='object')
```

In [ ]:

In [ ]: