In [1]:
```
import pandas as pd
covidData = pd.read_csv("Datasets/covid.csv")
```

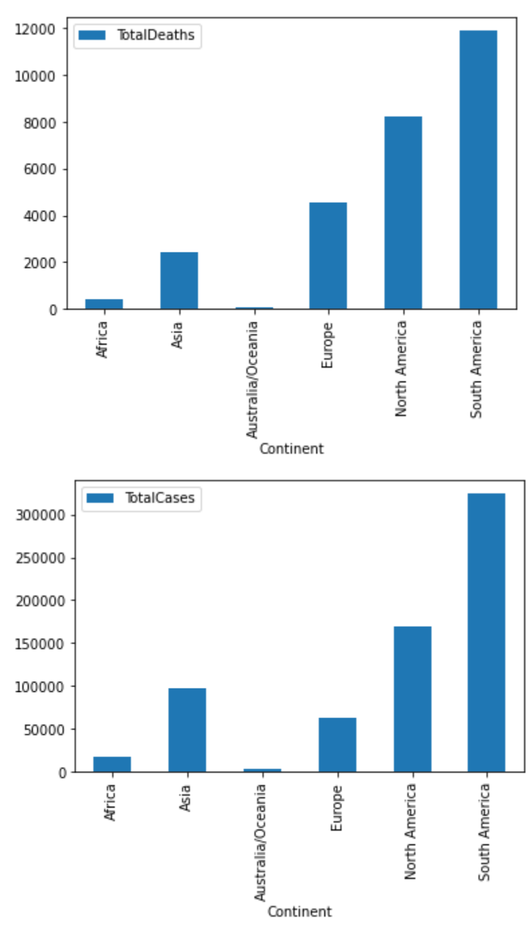
In [2]:
```
covidData.head()
```

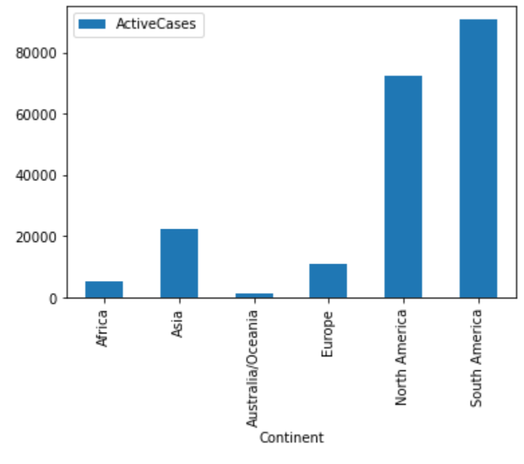
Out[2]:

	Country/Region	Continent	Population	TotalCases	NewCases	TotalDeaths	NewDeaths	TotalRecovered	NewRecovered	ActiveCases	Serious,Critical	Tot Cases/1M pop	Deaths/1M pop	TotalTests	Tests/1M pop	WHO Region
0	USA	North America	3.311981e+08	5032179	NaN	162804.0	NaN	2576668.0	NaN	2292707.0	18296.0	15194.0	492.0	63139605.0	190640.0	Americas
1	Brazil	South America	2.127107e+08	2917562	NaN	98644.0	NaN	2047660.0	NaN	771258.0	8318.0	13716.0	464.0	13206188.0	62085.0	Americas
2	India	Asia	1.381345e+09	2025409	NaN	41638.0	NaN	1377384.0	NaN	606387.0	8944.0	1466.0	30.0	22149351.0	16035.0	South-EastAsia
3	Russia	Europe	1.459409e+08	871894	NaN	14606.0	NaN	676357.0	NaN	180931.0	2300.0	5974.0	100.0	29716907.0	203623.0	Europe
4	South Africa	Africa	5.938157e+07	538184	NaN	9604.0	NaN	387316.0	NaN	141264.0	539.0	9063.0	162.0	3149807.0	53044.0	Africa

In [24]:
```
covidData.drop(['NewCases', 'Serious,Critical', 'Tot Cases/1M pop', 'Tests/1M pop', 'Deaths/1M pop', 'ActiveCases', 'NewRecovered', 'TotalRecovered', 'NewDeaths', 'TotalTests', 'TotalCases', 'Population'], axis=1).groupby("Contin
covidData.drop(['NewCases', 'Serious,Critical', 'Tot Cases/1M pop', 'Tests/1M pop', 'Deaths/1M pop', 'ActiveCases', 'NewRecovered', 'TotalRecovered', 'NewDeaths', 'TotalTests', 'TotalDeaths', 'Population'], axis=1).groupby("Conti
covidData.drop(['NewCases', 'Serious,Critical', 'Tot Cases/1M pop', 'Tests/1M pop', 'Deaths/1M pop', 'TotalCases', 'NewRecovered', 'TotalRecovered', 'NewDeaths', 'TotalTests', 'TotalDeaths', 'Population'], axis=1).groupby("Contin
```

Out[24]:
```
<AxesSubplot:xlabel='Continent'>
```





I compared ActiveCases, TotalDeaths and TotalCases with eachother grouped by continent to see the correlations. The 3 graphs all look kind off similar. It seems that they correlate with eachother. When a continent has alot of ActiveCases it also has alot of TotalCases and TotalDeaths.