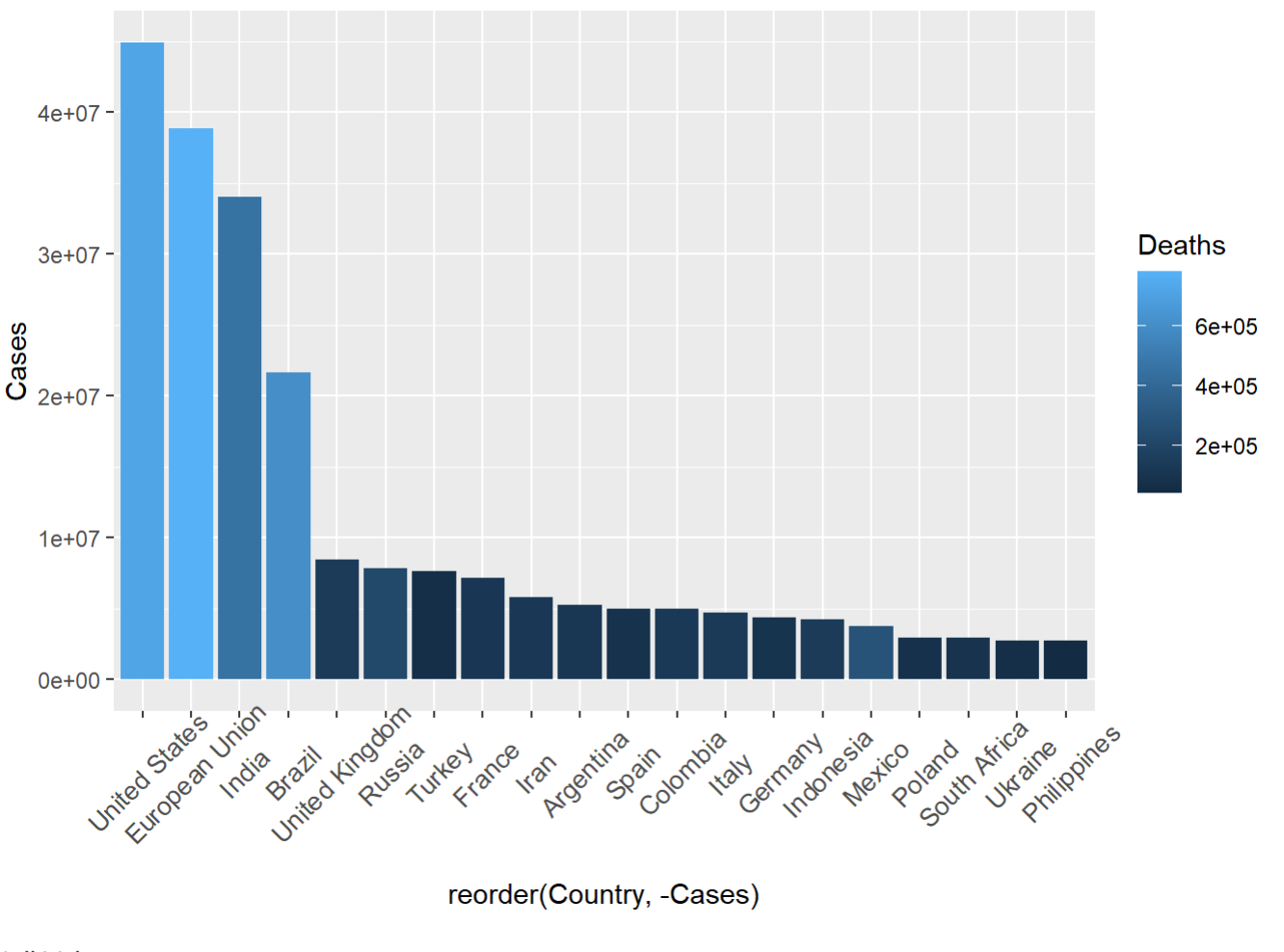
Pandemic Terminator: An analysis of pandemic impacts and responses

```r
#Section 2. Plotting and analysis.
country_data [-c(1),] %>%
  top_n(20) %>%
  ggplot(aes(x = reorder(Country,-Cases),  y= Cases, fill = Deaths ))+
  geom_col() +
  theme(axis.text.x=element_text(angle=45,size=10,hjust = 0.5,vjust=0.7))
```

```
## Selecting by Cases
```



Null Values:

```r
miss_var_summary(cases_pop)
```

```
## # A tibble: 10 x 3
##    variable          n_miss pct_miss
##    <chr>              <int>    <dbl>
##  1 Migrants (net)        22   11.2
##  2 Population (2020)     15    7.65
##  3 Yearly Change         15    7.65
##  4 Net Change            15    7.65
##  5 Density (P/Km²)       15    7.65
##  6 Land Area (Km²)       15    7.65
##  7 Deaths/million         8    4.08
##  8 Deaths                 5    2.55
##  9 Cases                  1    0.510
## 10 Country                0    0
```

```r
miss_case_summary(cases_pop)
```

```
## # A tibble: 196 x 3
##     case n_miss pct_miss
##    <int>  <int>    <dbl>
##  1   196      9       90
##  2   192      7       70
##  3   194      7       70
##  4     1      6       60
##  5     8      6       60
##  6    33      6       60
##  7    39      6       60
##  8    66      6       60
##  9    77      6       60
## 10   100      6       60
## # ... with 186 more rows
```

```
## # A tibble: 8 x 10
##   Country         `Deaths/million` Deaths Cases `Population (20~ `Yearly Change`
##   <chr>                      <dbl>  <dbl> <dbl>            <dbl>           <dbl>
## 1 "Samoa"                       NA     NA     3           198414            0.67
## 2 "Marshall Isla~               NA     NA     4            59190            0.68
## 3 "Solomon Islan~               NA     NA    20           686884            2.55
## 4 "Vatican City"                NA      0    27               NA           NA
## 5 "Palau"                       NA      0     8            18094            0.48
## 6 "Federated Sta~               NA      0     1               NA           NA
## 7 "Kiribati"                    NA     NA     2           119449            1.57
## 8 ".mw-parser-ou~               NA     NA    NA               NA           NA
## # ... with 4 more variables: Net Change <dbl>, Density (P/Km²) <dbl>,
## #   Land Area (Km²) <dbl>, Migrants (net) <dbl>
```

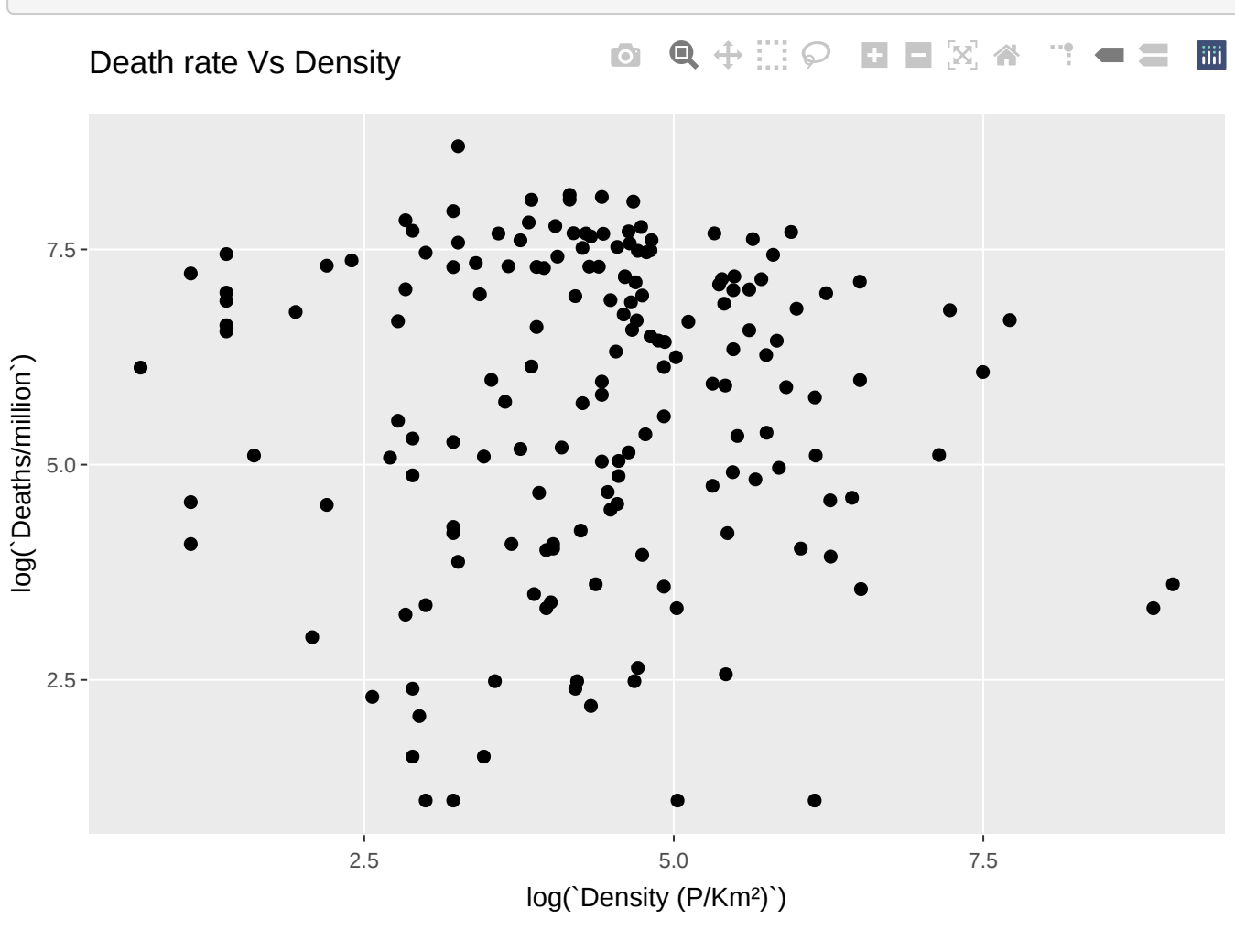
```
## Warning: Unknown or uninitialised column: `pop`.
```

```
## NULL
```

Modeling Impacts of Covid 19:

```r
death_density <- ggplot(cases_pop, aes(x=log(`Density (P/Km²)`), y=log(`Deaths/million`),
                                       text = Country)) +
  geom_point() +
  ggtitle("Death rate Vs Density")

ggplotly(death_density)
```



Regression, different variables to explain deaths.

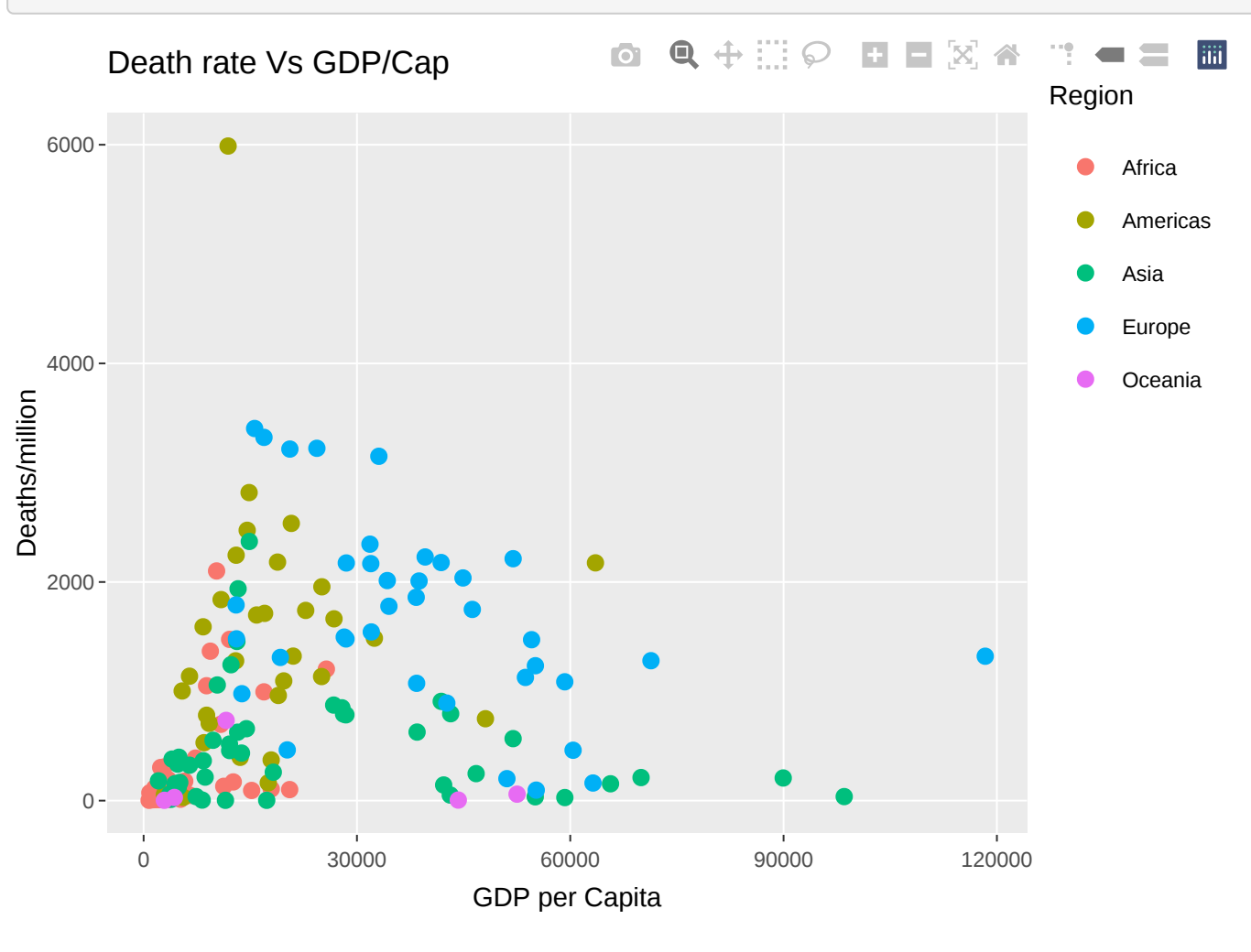
```r
##Regression: How the number of cases, density, population and migrants affect the death rate.
death_density_reg <- lm(log(`Deaths/million`) ~  log(`Density (P/Km²)`)
                        +log(GDPperCap),
                        data = cases_pop_gdp)
summary(death_density_reg)
```

```
##
## Call:
## lm(formula = log(`Deaths/million`) ~ log(`Density (P/Km²)`) +
##     log(GDPperCap), data = cases_pop_gdp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.3678 -0.8716  0.3501  1.0144  2.9047
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)              -2.15256    1.00462  -2.143   0.0336 *
## log(`Density (P/Km²)`) -0.08583    0.08337  -1.029   0.3048
## log(GDPperCap)            0.87663    0.10365   8.457 1.43e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.543 on 164 degrees of freedom
## Multiple R-squared:  0.3037, Adjusted R-squared:  0.2952
## F-statistic: 35.76 on 2 and 164 DF,  p-value: 1.288e-13
```

Q.1: How has GDP/Density impacted deaths across different regions.

In order to determine the extent of the detrimental effects of COVID-19, a metric must be chosen to represent/summarize the degree to which a country has been impacted by COVID-19. In the case of this analysis, the decided upon metric that most faithfully achieves this was determined to be Deaths per capita. Having established this the "Death rate Vs Density" scatter plot was constructed. Upon visual inspection of this graph, it can be interpreted that there is no clear relationship between deaths per capita and population density (measured in density per Km). In order to strengthen this contention, a linear regression was also plotted which included variables log(GDP per capita), log(density). This regression supports the notion that there evidence of no relation as the p-value of the density regressor implies it is not statistically significant.

```r
#cases_pop_gdp$`Extimate.x`<-as.numeric(cases_pop_gdp$`Extimate.x`)
anim <- ggplot(cases_pop_gdp, aes(x= GDPperCap,
                                  y=`Deaths/million`,
                                  text = Country)) +
  geom_point(aes(colour = Region), size = 2)+
  labs(x='GDP per Capita')+
  ggtitle("Death rate Vs GDP/Cap")
ggplotly(anim)
```



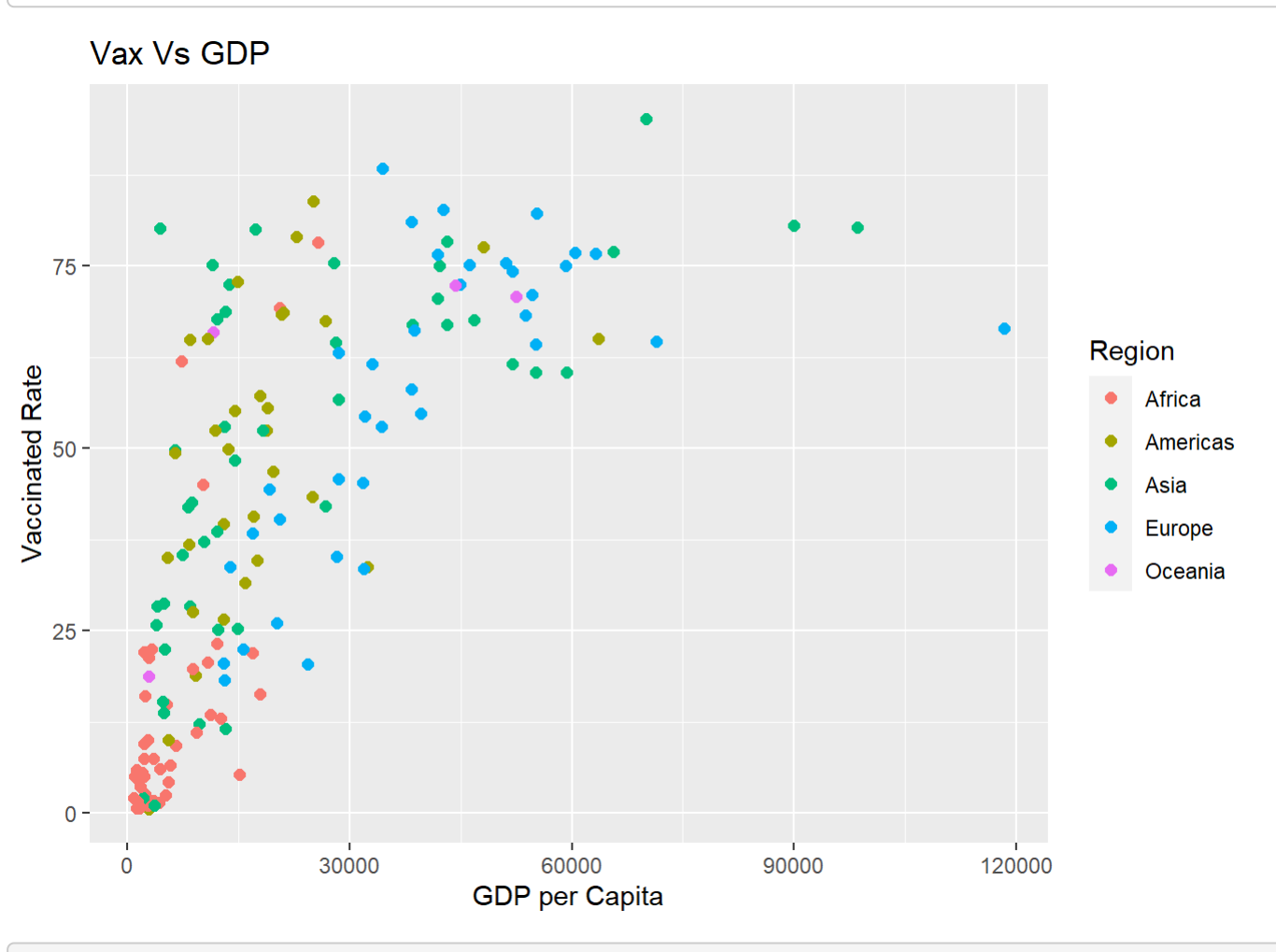
Conversely to population density. GDP per capita does appear to share a positive relationship with deaths per million (from graph "Death rate Vs GDP/Cap") . That is the higher the GDP of a country the higher the death rate. This conclusion, continues if we look back the the regression, a p value of 1.68e-14 is given, suggesting little evidence against the null hypothesis that GDP and death rates are not correlated.

from graph "Death rate Vs GDP/Cap", it would appear that Europe is heavily affected by Covid, where as the region of Africa and most Asian regions (with lower GDP's) also have very low Covid death rates. It can also be seen that The Americas are highly impacted by Covid as well.

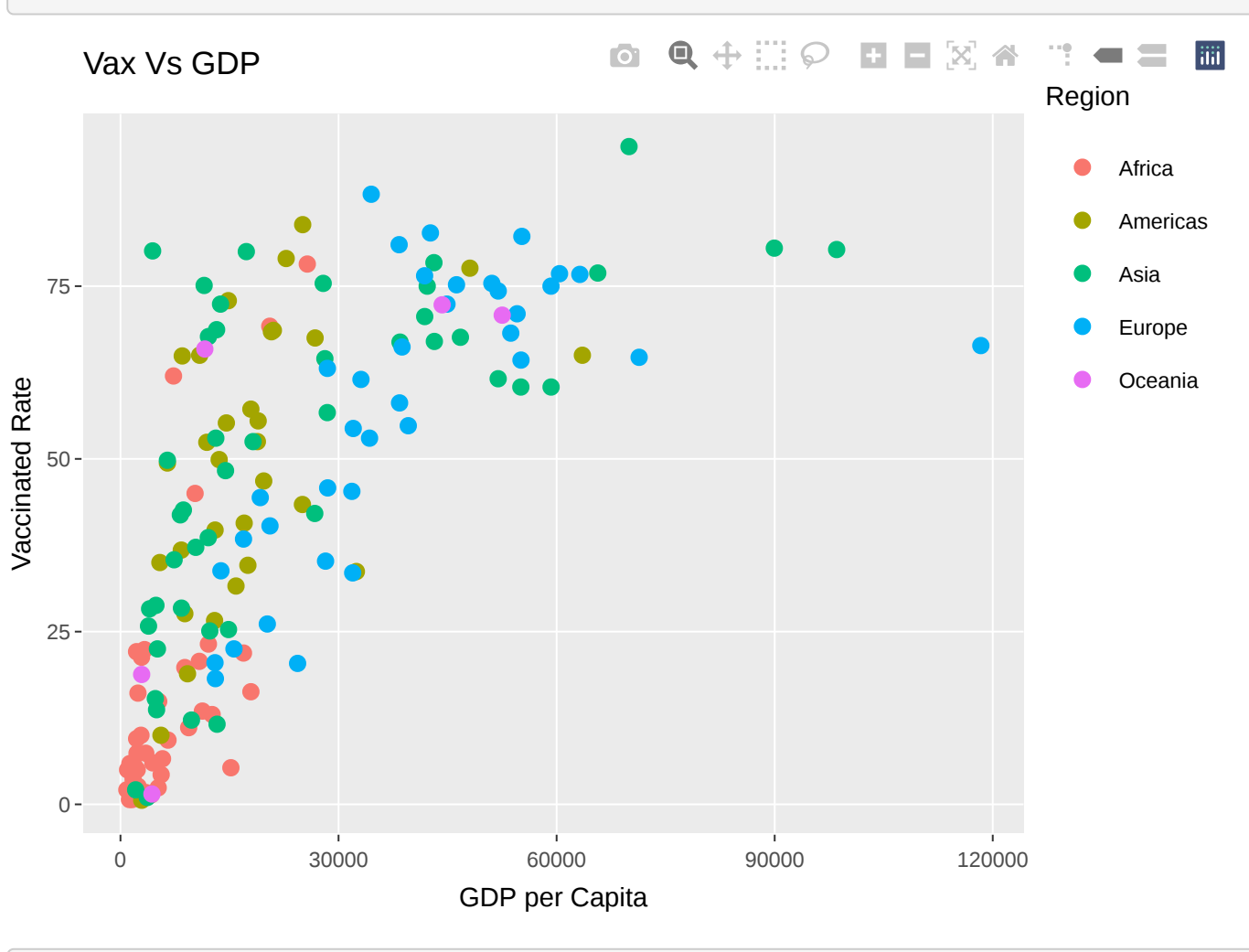
Explaining Vaccination rates:

```r
##How the GDP per capita affects the vaccinate rate.
cases_pop_gdp_vax %>%
  ggplot(aes(x=GDPperCap,
             y=Percentage,
             text = Country)) +
  geom_point(aes(colour = Region), size = 2) +
  labs(x='GDP per Capita',y = "Vaccinated Rate")+
  ggtitle("Vax Vs GDP")
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```

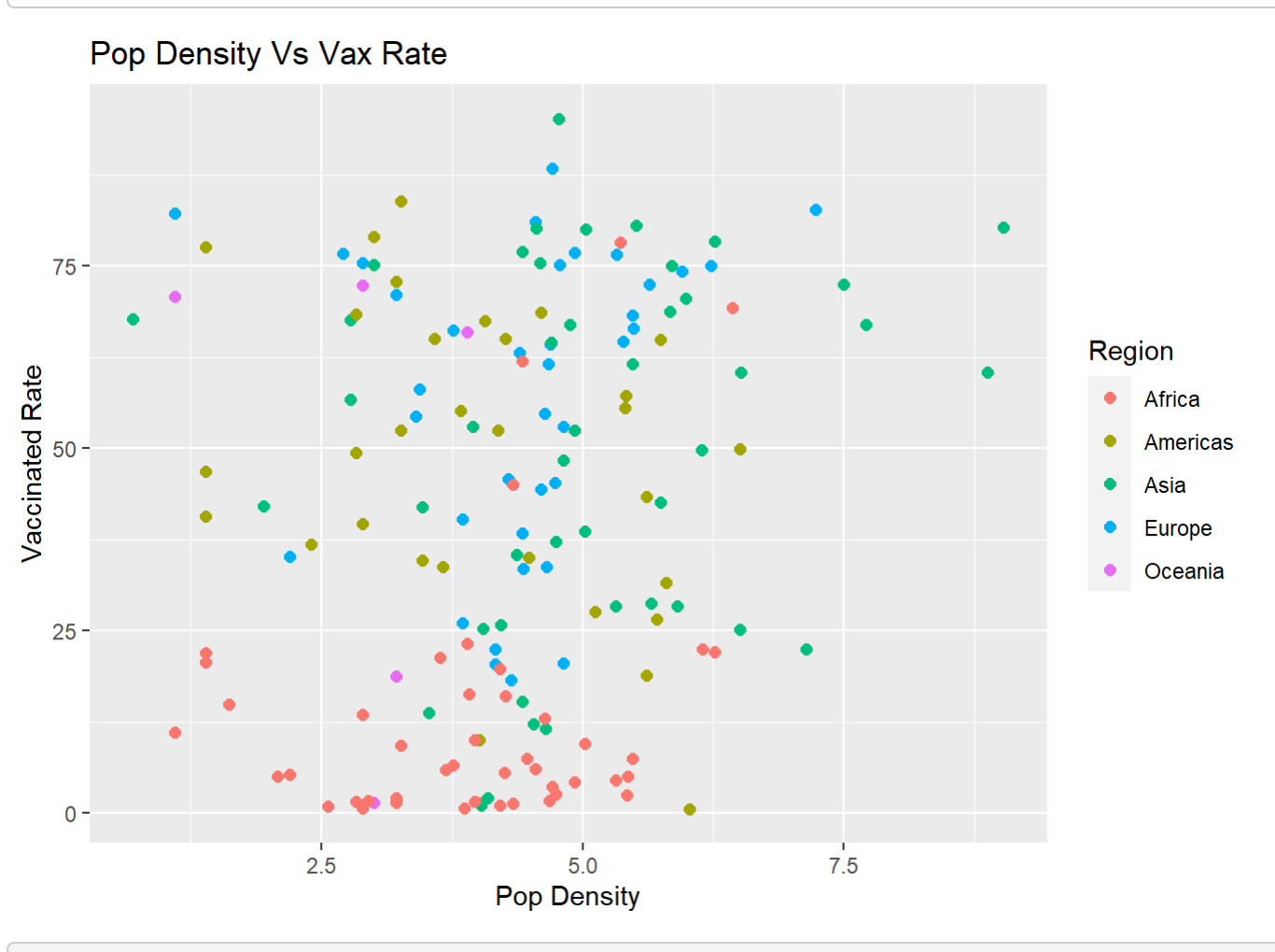



```r
ggplotly()
```

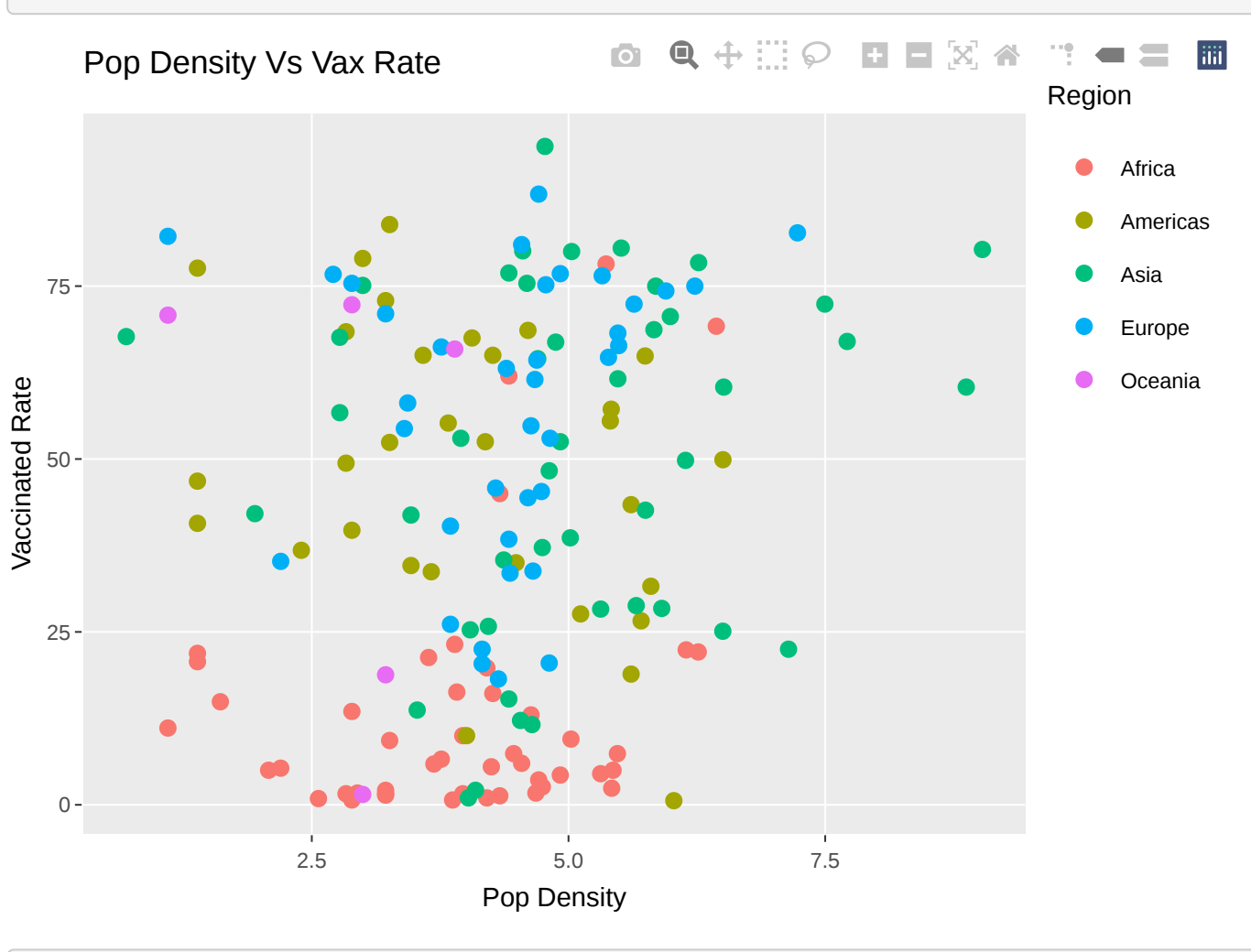



```r
cases_pop_gdp_vax %>%
  ggplot(aes(x=log(`Density (P/Km²)`),
             y=Percentage,
             text = Country)) +
  geom_point(aes(colour = Region), size = 2) +
  labs(x='Pop Density',y = "Vaccinated Rate")+
  ggtitle("Pop Density Vs Vax Rate")
```

```
## Warning: Removed 2 rows containing missing values (geom_point).
```




```r
ggplotly()
```




```r
vaccine_reg <- lm(Percentage ~  log(Cases) + log(`Density (P/Km²)`) + log(`Population (2020)`)
                  + log(GDPperCap)+log(`Deaths/million`),
                  data = cases_pop_gdp_vax)
summary(vaccine_reg)
```

```
##
## Call:
## lm(formula = Percentage ~ log(Cases) + log(`Density (P/Km²)`) +
##     log(`Population (2020)`) + log(GDPperCap) + log(`Deaths/million`),
##     data = cases_pop_gdp_vax)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -39.290 -10.593   0.437   9.476  59.390
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -106.5325    25.9667  -4.103 6.51e-05 ***
## log(Cases)                   4.7268     1.8997   2.488   0.0139 *
## log(`Density (P/Km²)`)       0.9984     0.8716   1.146   0.2537
## log(`Population (2020)`)    -3.8439     1.8477  -2.080   0.0391 *
## log(GDPperCap)              18.0849     1.5011  12.048  < 2e-16 ***
## log(`Deaths/million`)       -4.0293     1.6677  -2.416   0.0168 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.92 on 159 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.6786, Adjusted R-squared:  0.6685
## F-statistic: 67.16 on 5 and 159 DF,  p-value: < 2.2e-16
```

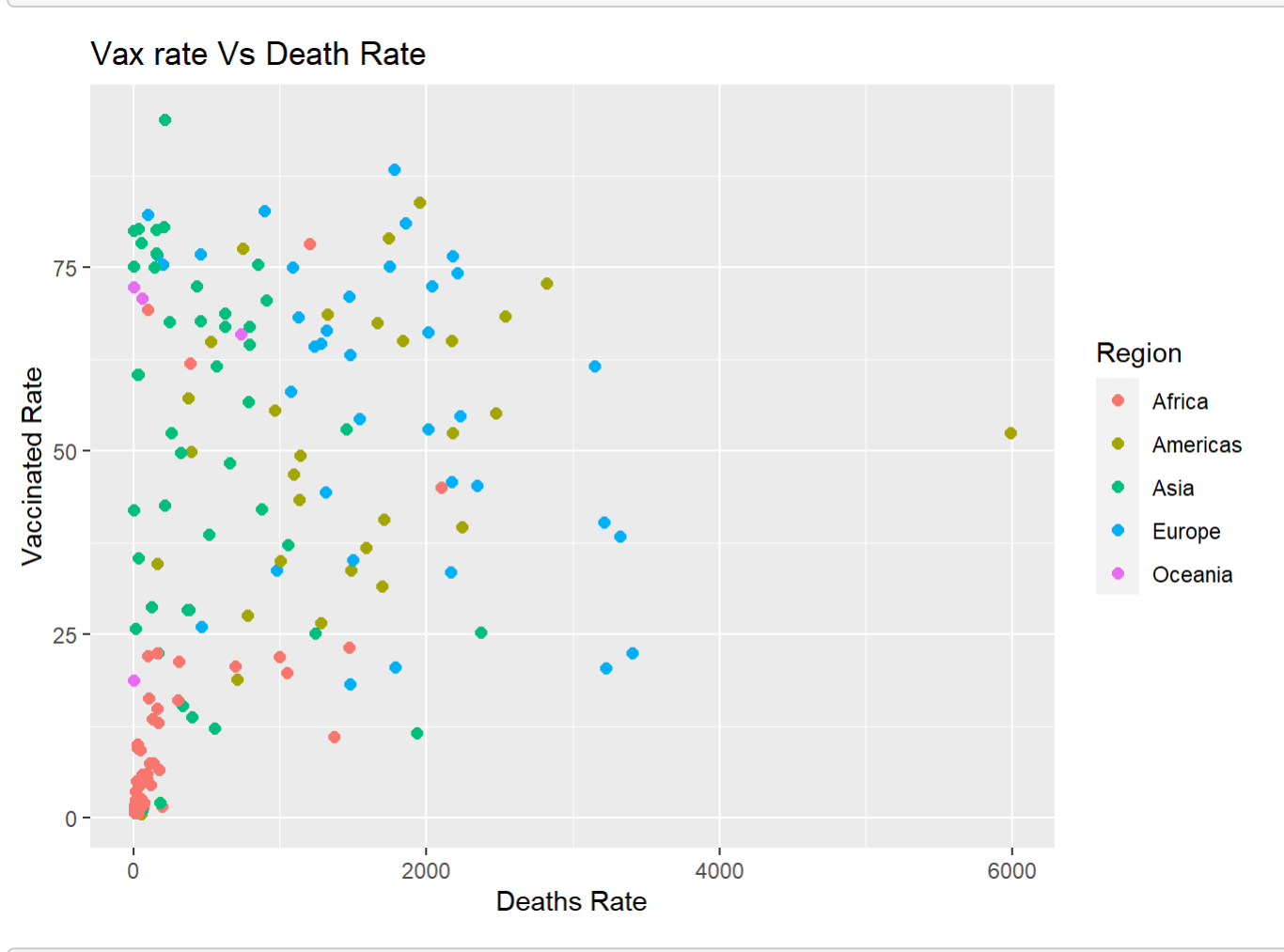
Q.2: Will the population density impact on the vaccination rates?

In a similar style of analysis to the previous question in order to review the impacts of GDP per capita and population density upon vaccination rate, both scatter plot and regression analysis was conducted. Beginning with population density it can initially be inferred from the scatter plot, "pop density vs vax rate", that there is no strong correlation between the two variables, although it can be seen that African countries have a relatively low vaccination rate. Such a hypothesis is confirmed in the regression was we are able to see that population density is not statistically significant in predicting vaccination rate (p-value < 0.05)

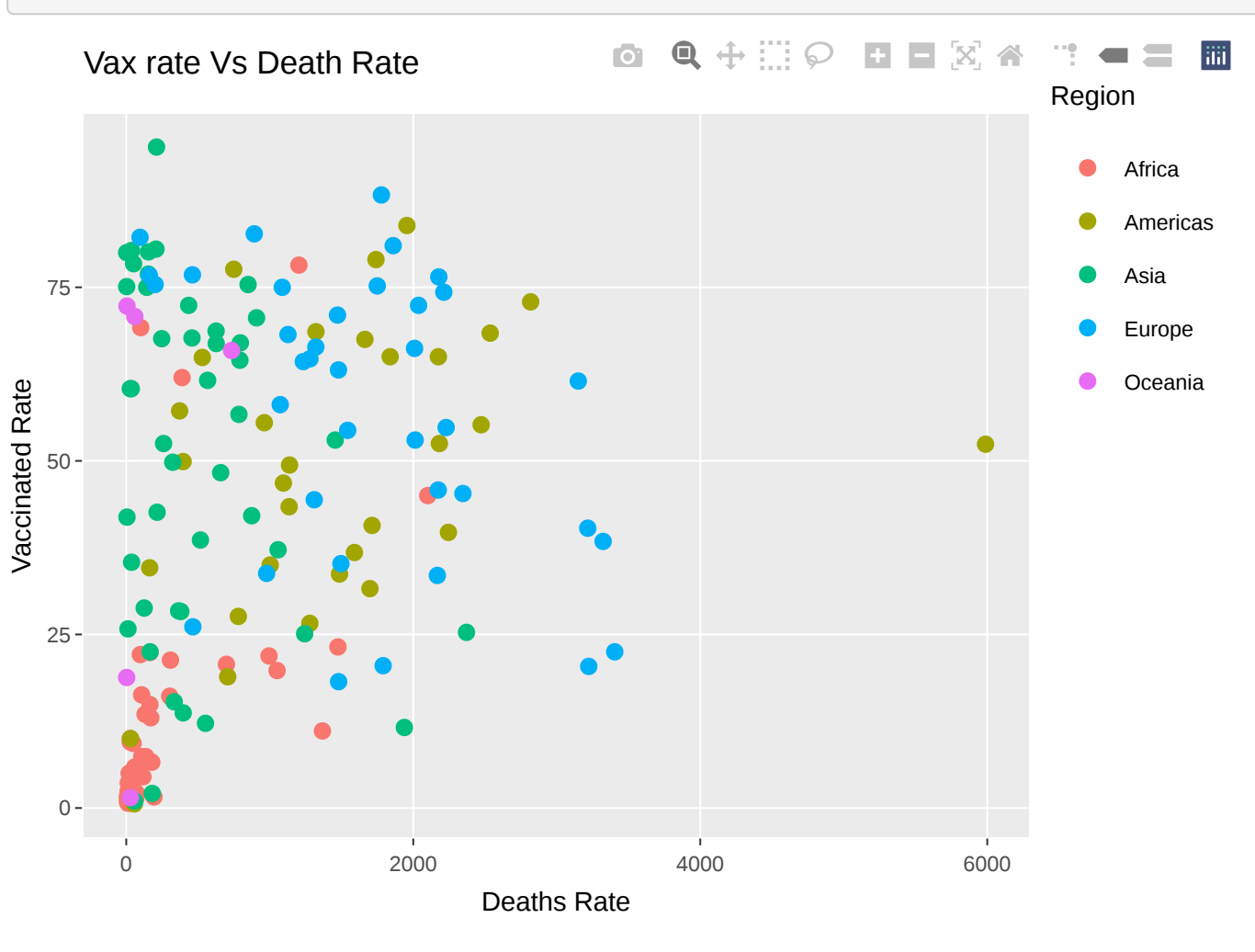
Q.4: Are the vaccination roll outs across the world, bias to richer countries?

From the regression above, The GDP per capita appears significant, which means the GDP has great impact on the vaccine rate. This can be seen from its p value, suggesting low evidence of the correlation of these two variables being 0. The GDP per capital increases 1% on average, the vaccine rate will increase 18.07%, holding other variables constant. Using the graph, "Vax rate Vs GDP" above, it can be seen clearly a strong positive relationship between the vaccination rates and GDP. This gives us an indication that vaccination rates are severely bias, and favored towards richer countries. This can be extrapolated to regions as well. From the graphic, the continent Africa can be seen, due to its lower GDP, has its countries distributed among very vaccination rates, where as the richer countries on average, Europe appear to be highly distributed with higher vaccination rates.

```r
cases_pop_gdp_vax %>%
  ggplot(aes(x=`Deaths/million`,
             y=Percentage,
             text = Country)) +
  geom_point(aes(colour = Region), size = 2) +
  labs(x='Deaths Rate',y = "Vaccinated Rate") + ggtitle("Vax rate Vs Death Rate")
```




```r
ggplotly()
```



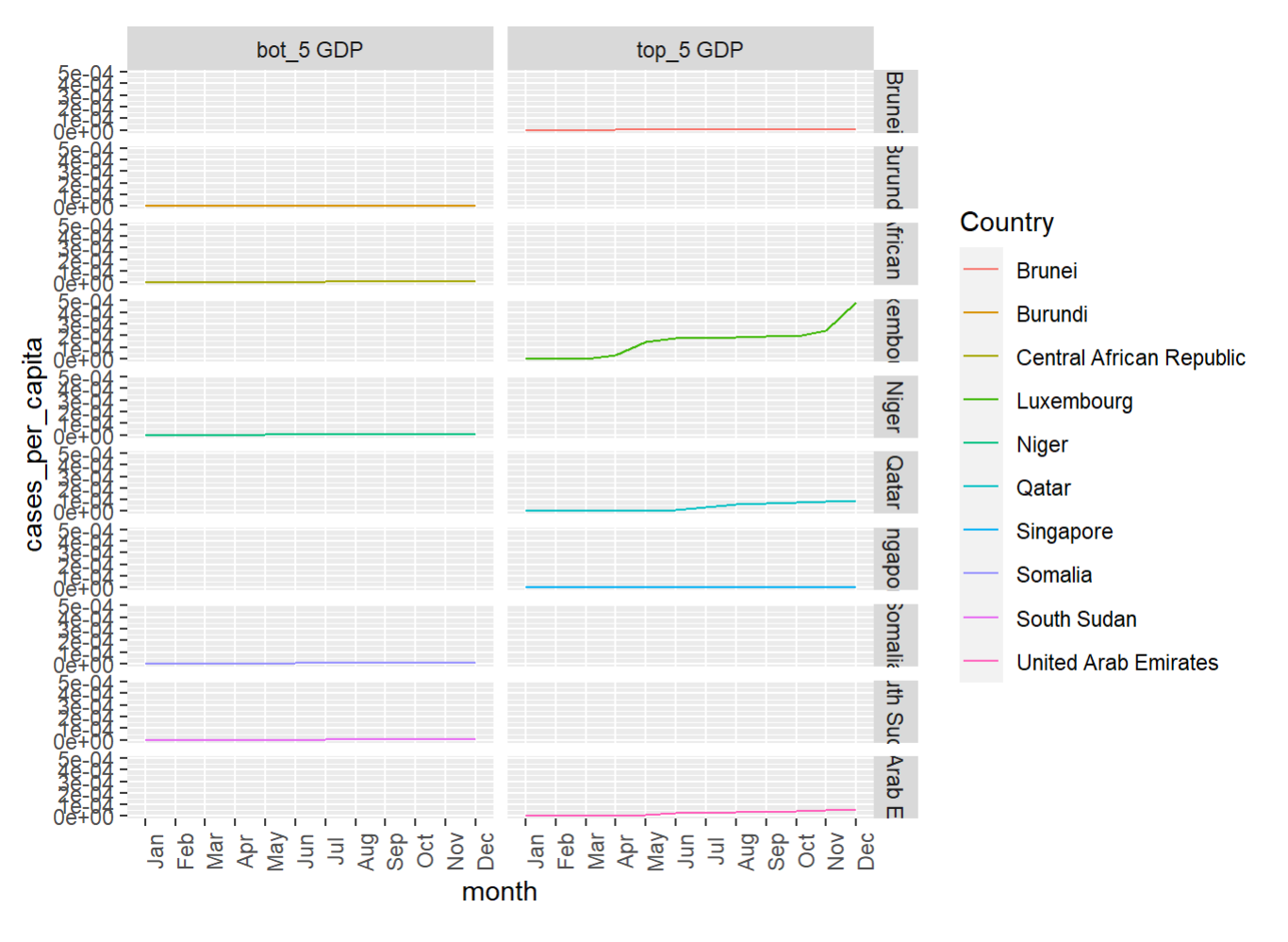
From the graph above, we can answer the question if countries impacted by covid have a better vaccination response. Qualitatively from this graphic above, the death rate does appear to have some positive correlation, Further more, the p value of 0.016 from the regression suggests there is little evidence of deaths and vaccination rates, not being correlated (coefficient of 0), reiterating the point that countries worse hit by covid appear to have smoother vaccination roll outs.

Time Series Analysis:

This section will look at how GDP, and vaccinations for the top and bottom countries from these two metrics change over time.

Creating the bottom and top 5 GDP for countries and Vaccination rates

```r
#combine the GDP per Capital top_5 country and bottom_5 country
comb_gdp <- bind_rows(y2020_gdp_top5,y2020_gdp_bot5)
```

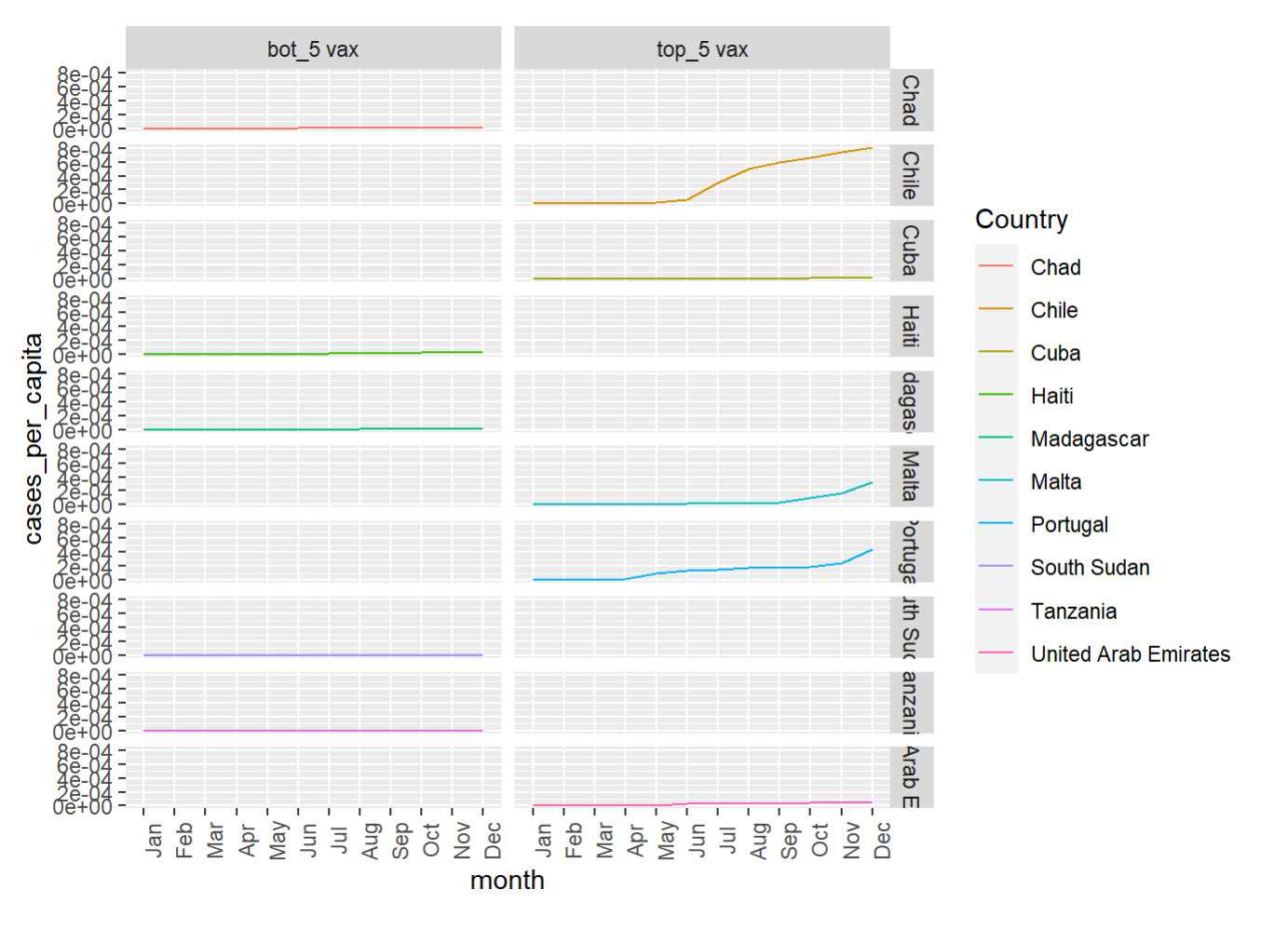


Q.4a: How does cases effect the bottom and top 5 countries based on GDP:

We picked five countries with high GDP, and five countries with low GDP. From the corresponding time series chart, it can be seen that countries with high GDP are more affected by COVID-19. The number of cases in most countries shows an increasing trend, while the curve in countries with low GDP is relatively flat, indicating that countries with high GDP are at higher risk.

Time Series Cases, for bottom and top 5 vaccination rates

```r
#graph data
comb_vacc %>%
  group_by(month) %>%
  ggplot(
      aes(x=month, y = cases_per_capita,group = 1,color = Country))+
  geom_line()+
  facet_grid(Country~group)+
  theme(axis.text.x = element_text(angle = 90))
```



Q.4b: How does cases effect the bottom and top 5 countries based on vaccine rates:

Further, we identified five countries with high and five with low vaccination rates. As you can see from the chart, countries with high vaccination rates show a higher trend of increasing cases, which further suggests that increasing cases drive vaccination rates in these countries. The fact that countries continue to be impacted by covid despite high vaccine rates, could come down to the delayed response vaccines will have, to improve countries' position in relation to covid. Potentially as time goes on, these trends could reverse.

Conclusion:

When analyzing the impacts of Covid across the world, fundamentally it appears richer countries have been impacted the worst in terms of cases and deaths. Wether this comes down to a lack of reporting from poorer countries, or has something to do with the viral characteristics of Covid itself, is a question still to be answered. From the vaccination rates, countries that are richer, also appear to have much more successful vaccination rates. This could be a combination of factors including the fact that not only have the resources to roll out vaccines, need it more as these countries have been disproportionally impacted by Covid-19. Finally, while this vaccination roll out is much more successful in rich countries, they still continue to be effected by Covid, potentially coming down to the delayed effect vaccinations will have on covid deaths.