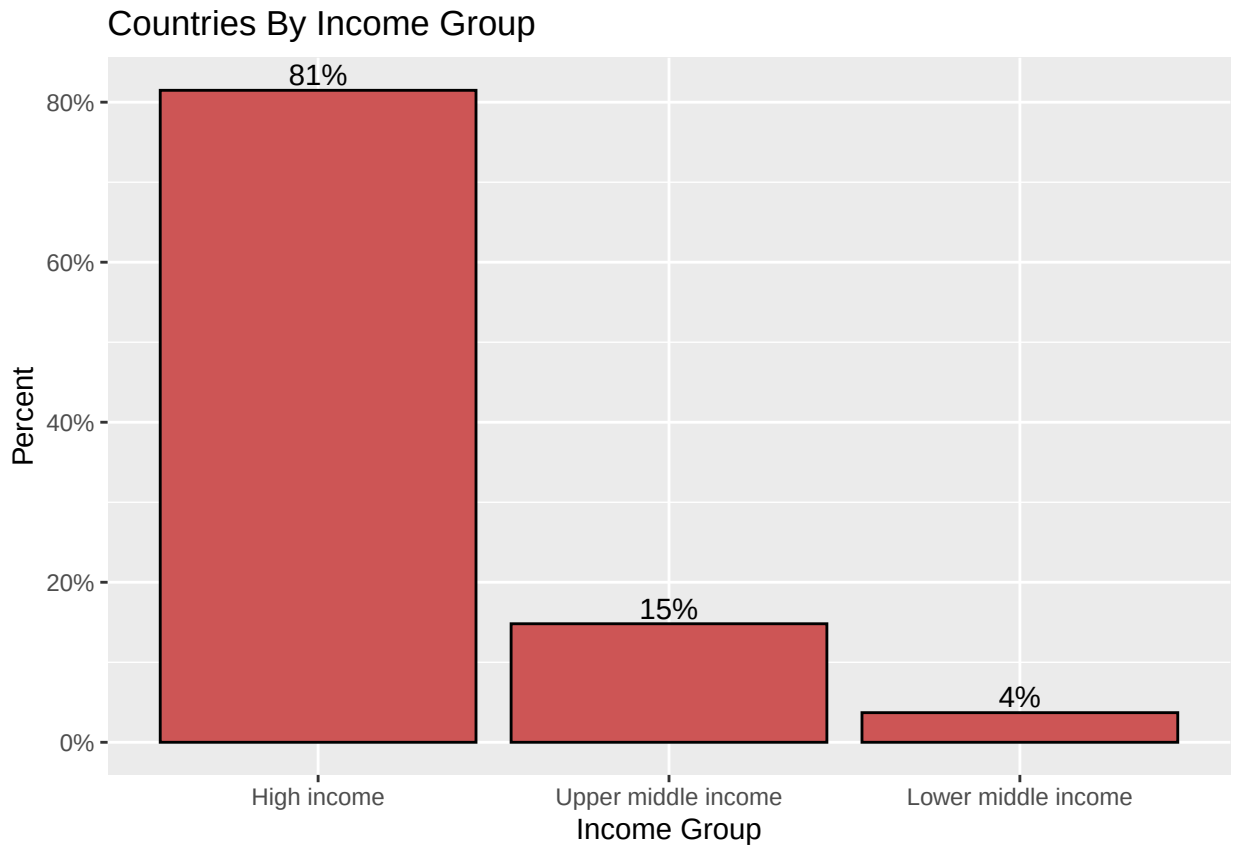
AI Vibrancy EDA

Taylor Boeckman

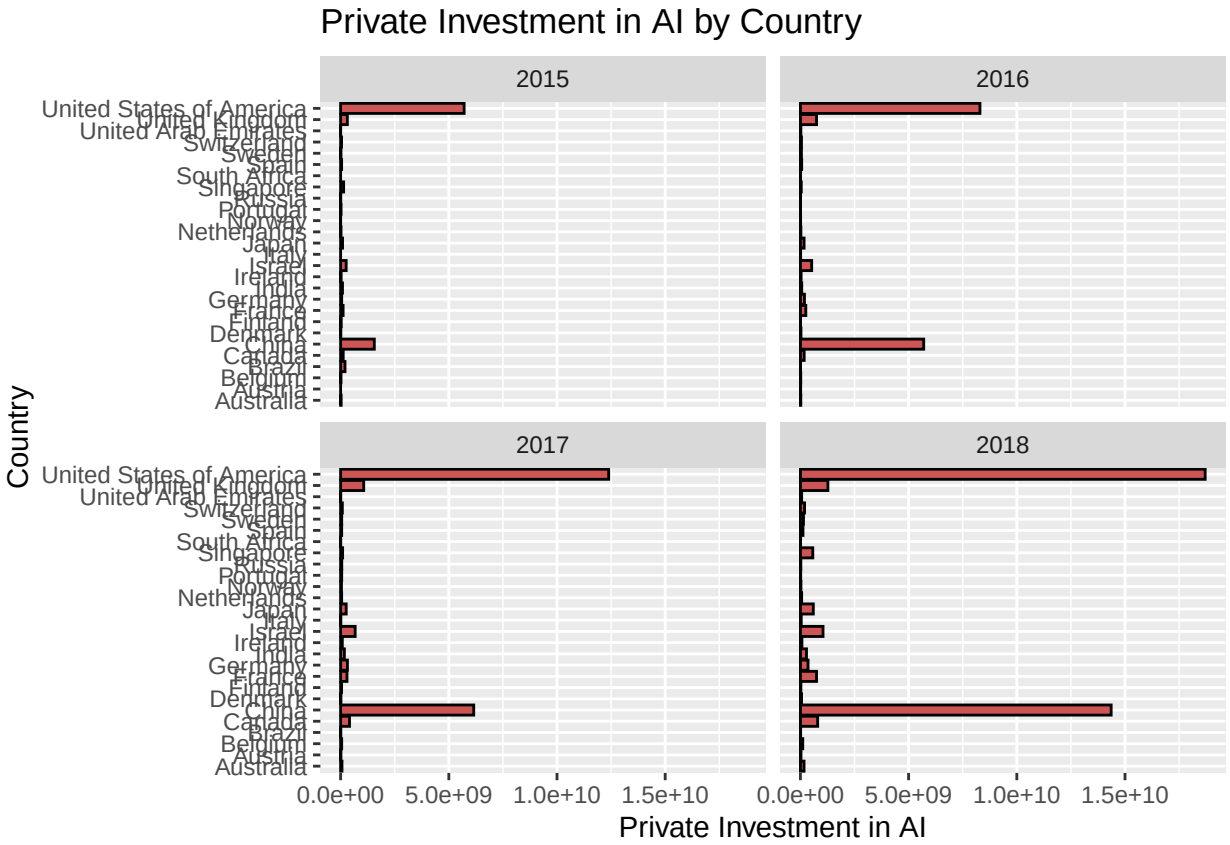
4/22/2020

```
aiVibrancyIndicators2019 = read.csv("2019indicators.csv")

aiVibrancyIndicators2019.df <- aiVibrancyIndicators2019

aiVibrancyIndicators2019.tib <- as_tibble(aiVibrancyIndicators2019.df)
```

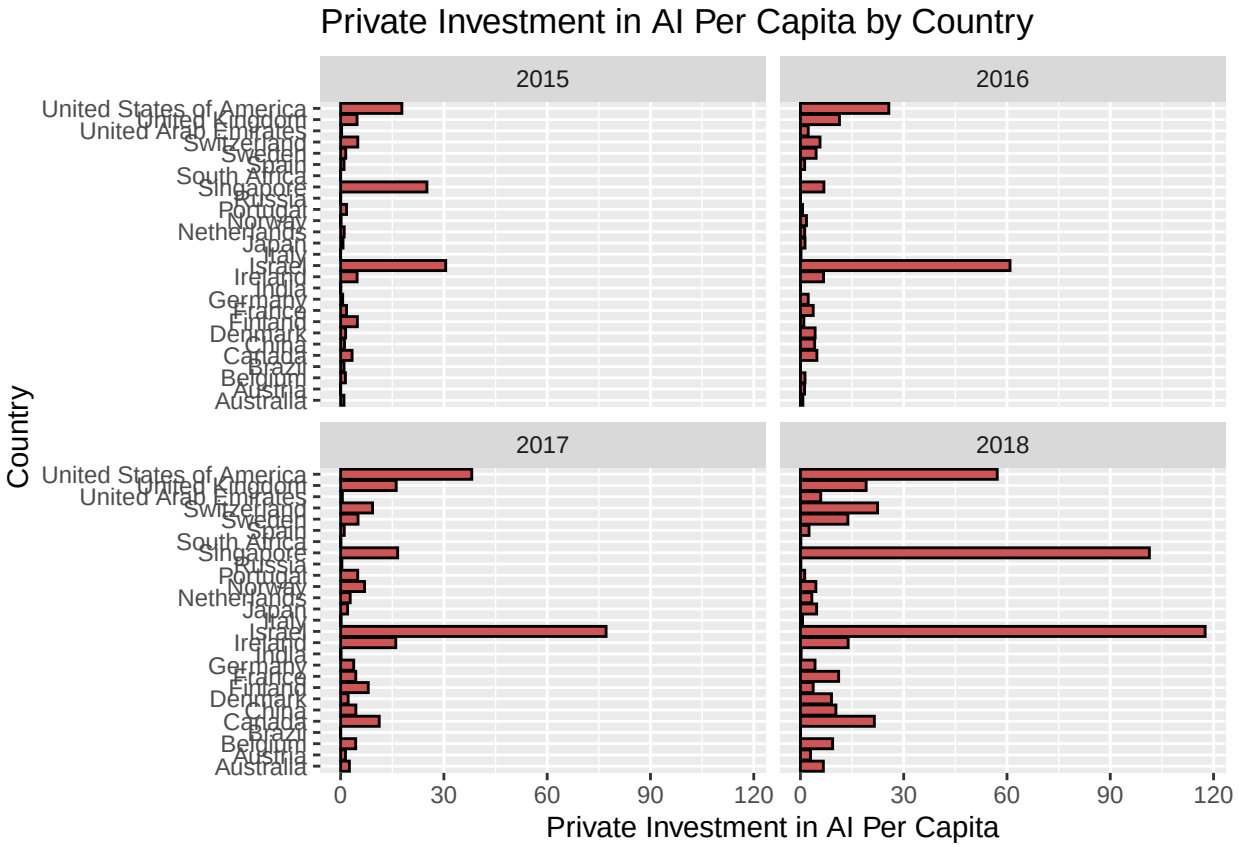
```
aiVibrancyIndicators2019.tib %>%
  count( income.group ) %>%
  mutate( pct = n / sum(n),
          pctlabel = paste0( round( pct*100 ), "%" ) ) %>%
  ggplot( aes( x = reorder( income.group, -pct ),
               y = pct ) ) +
      geom_col( fill = "indianred3",
                color = "black" ) +
      geom_text( aes( label = pctlabel ),
                      vjust = -0.25 ) +
      scale_y_continuous( labels = scales::percent ) +
      labs( x = "Income Group",
            y = "Percent",
            title = "Countries By Income Group" )
```



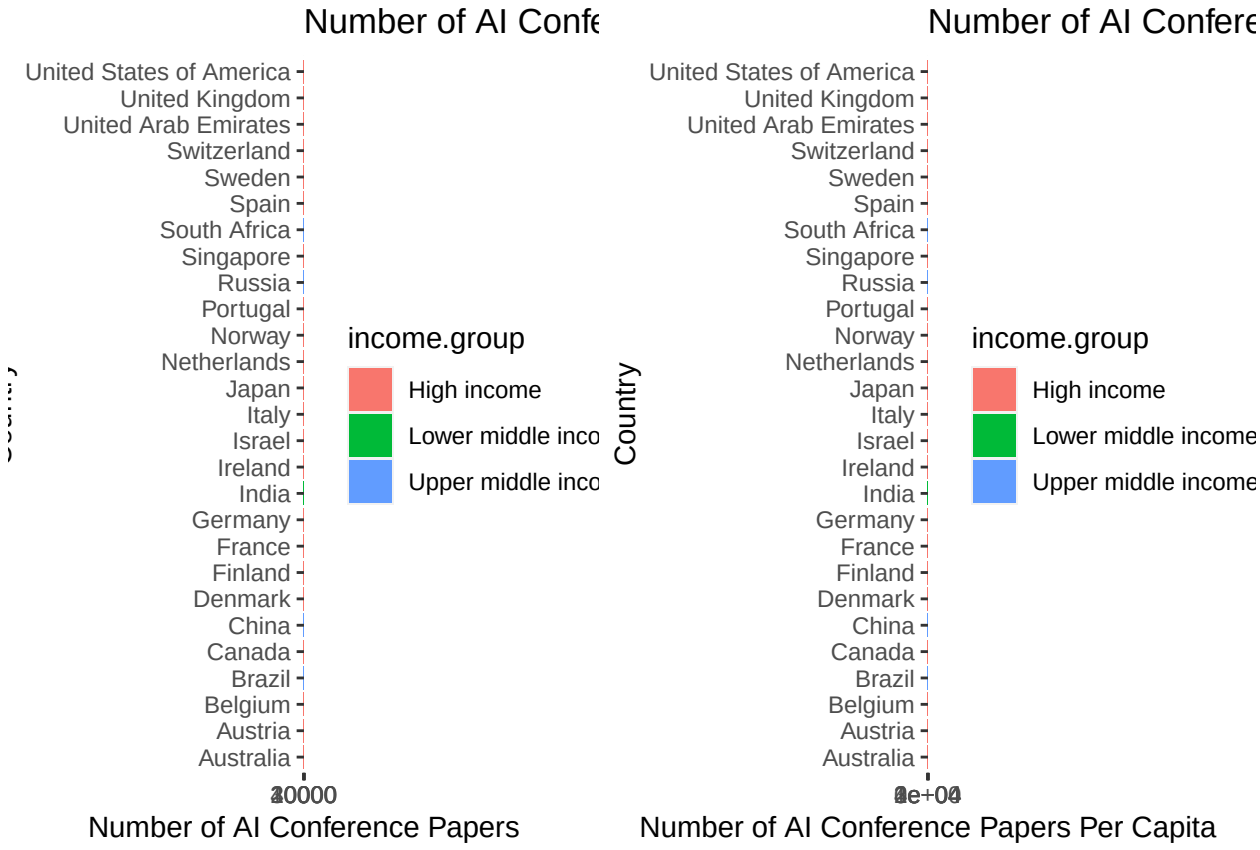

```
aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = ai_privateinvestment_tot , y = country ) ) +
      geom_col( fill = "indianred3",
                color = "black" ) +
      facet_wrap( ~ year, ncol = 2) +
      labs( x = "Private Investment in AI",
            y = "Country",
            title = "Private Investment in AI by Country" )
```



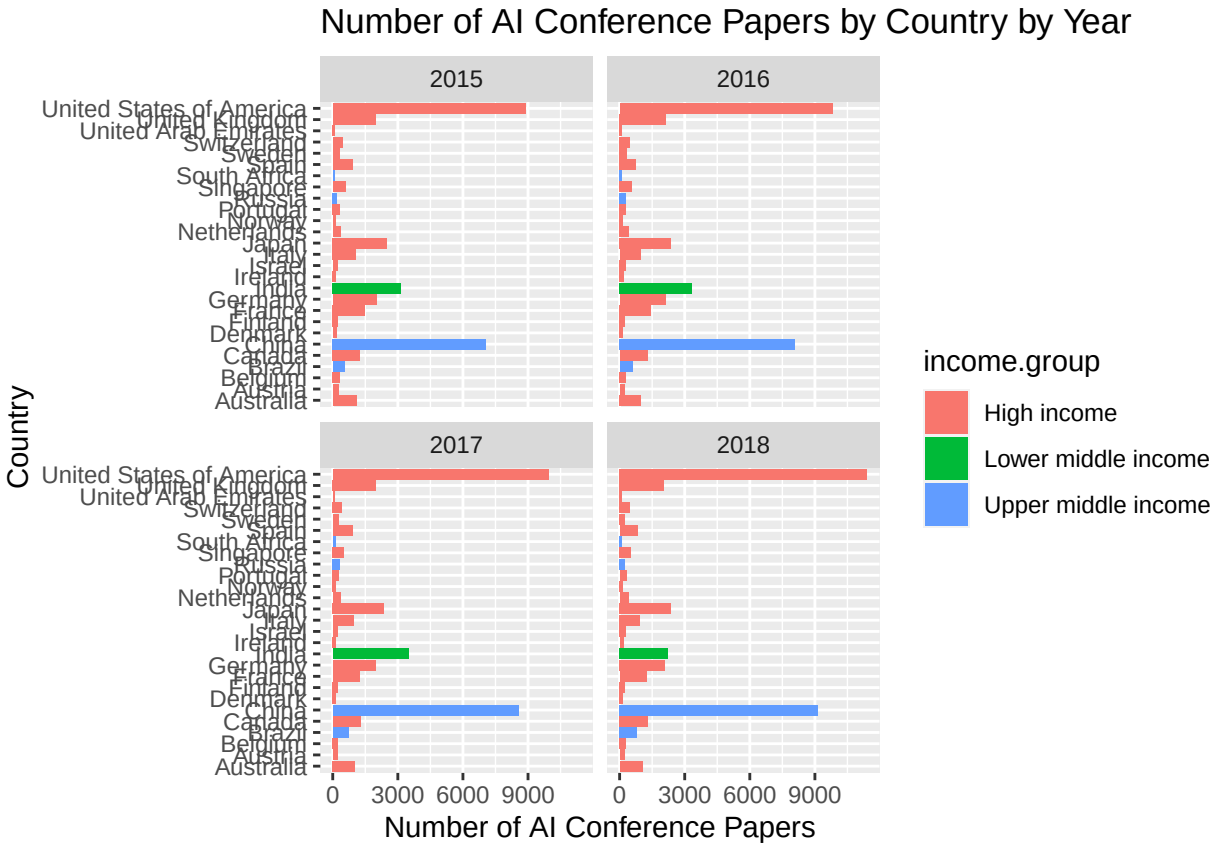

```
aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = ai_privateinvestment_tot_pc , y = country ) ) +
      geom_col( fill = "indianred3",
                color = "black" ) +
      facet_wrap( ~ year, ncol = 2) +
      labs( x = "Private Investment in AI Per Capita",
            y = "Country",
            title = "Private Investment in AI Per Capita by Country" )
```



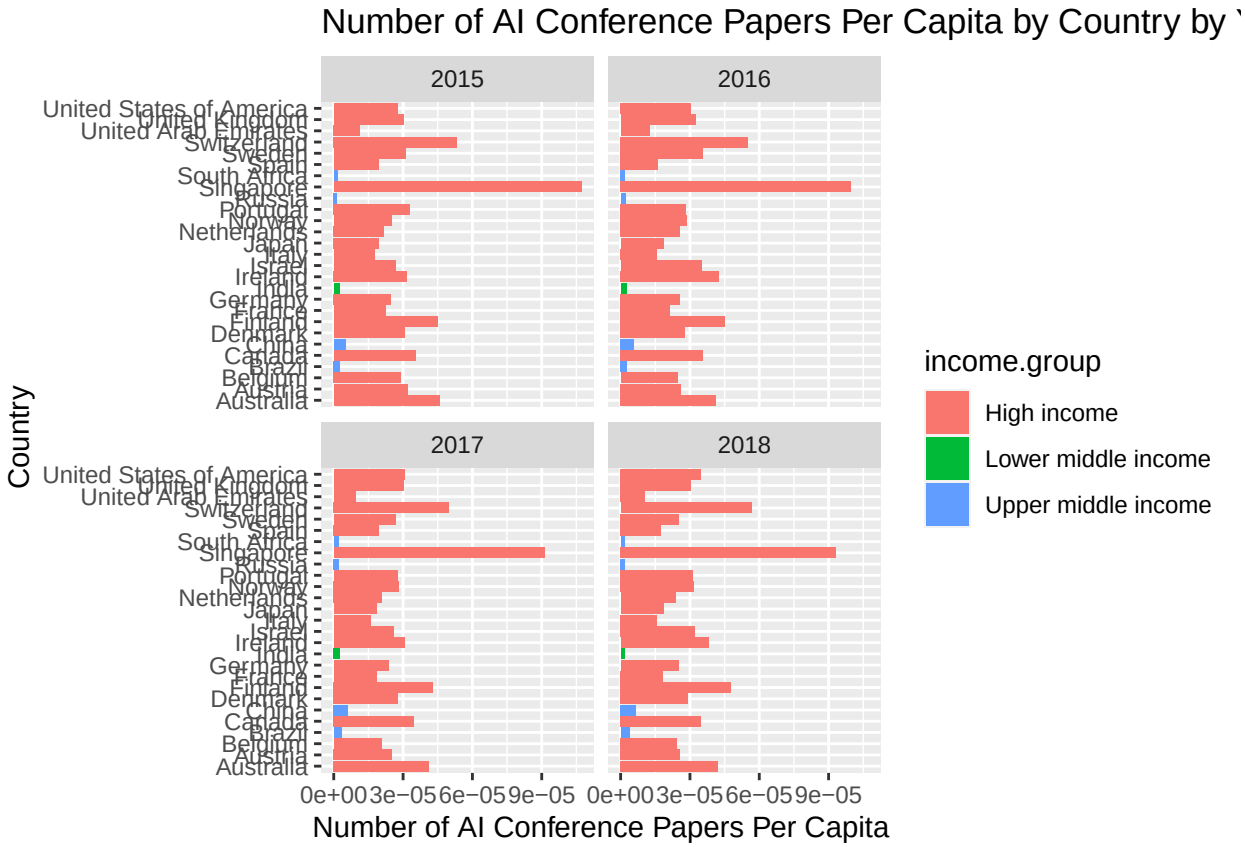

```
num_AIconf_papers_graph <- aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_papers, y = country, fill = income.group ) ) +
      geom_col() +
      labs( x = "Number of AI Conference Papers",
            y = "Country",
            title = "Number of AI Conference Papers by Country" )

num_AIconf_papers_pc_graph <- aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_papers_pc, y = country, fill = income.group ) ) +
      geom_col() +
      labs( x = "Number of AI Conference Papers Per Capita",
            y = "Country",
            title = "Number of AI Conference Papers  Per Capita by Country" )

grid.arrange(num_AIconf_papers_graph, num_AIconf_papers_pc_graph, ncol = 2)
```



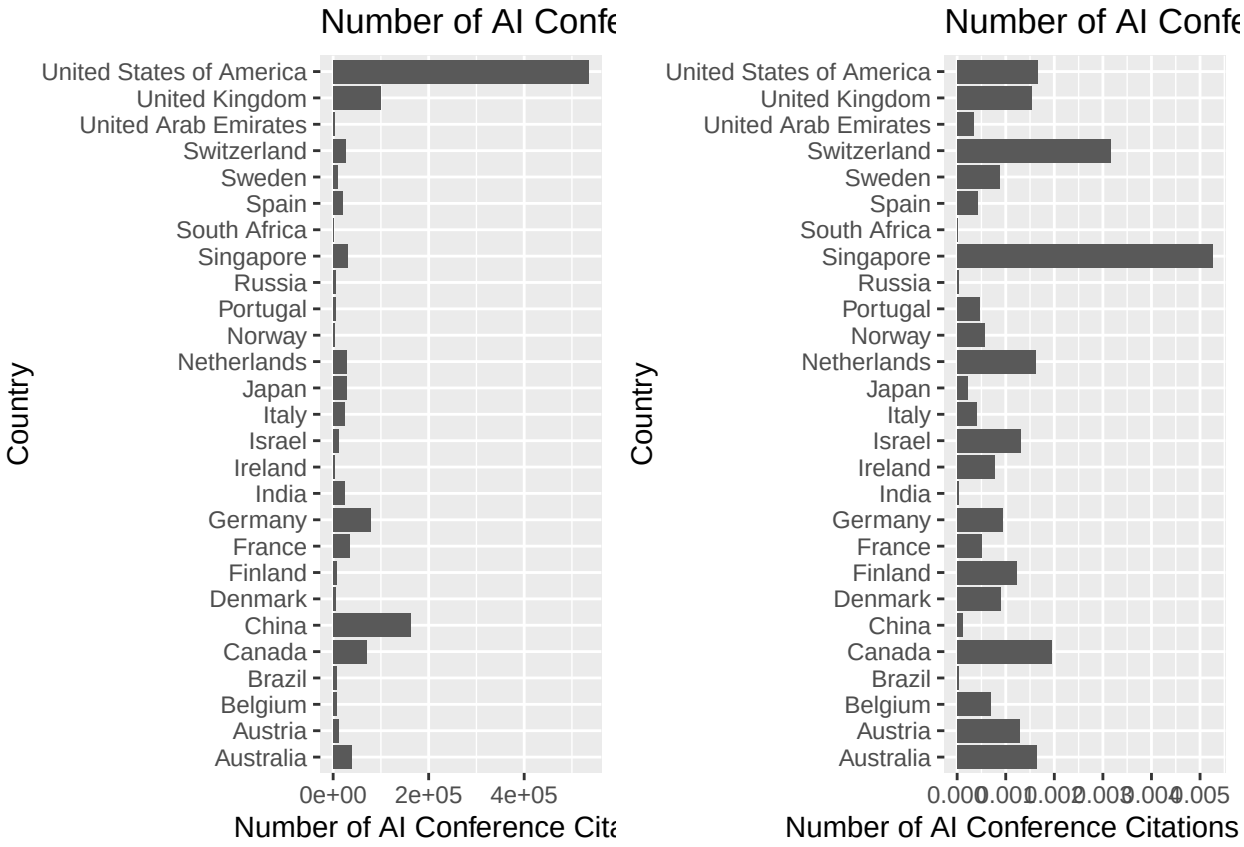

```
aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_papers, y = country, fill = income.group ) ) +
      geom_col() +
      facet_wrap( ~ year, ncol = 2 ) +
      labs( x = "Number of AI Conference Papers",
            y = "Country",
            title = "Number of AI Conference Papers by Country by Year" )
```



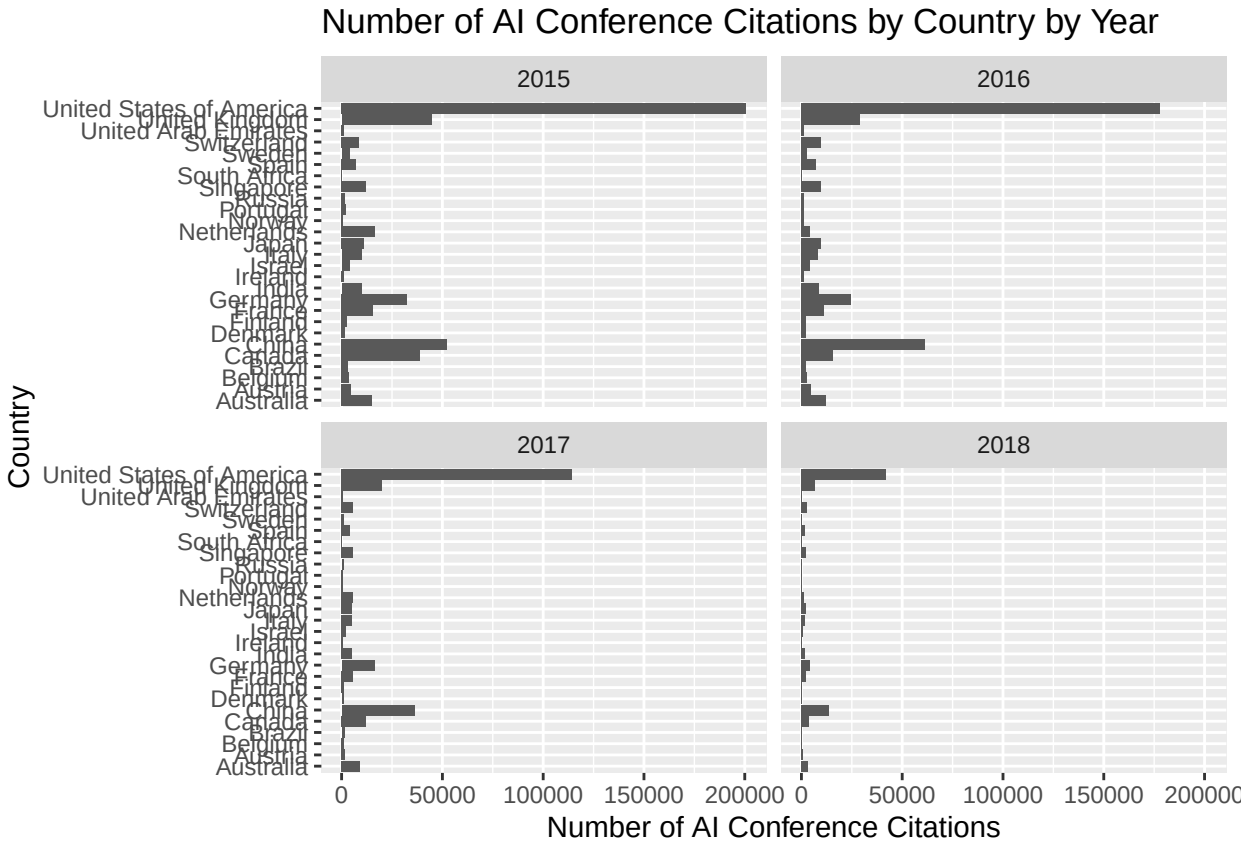

```
aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_papers_pc, y = country, fill = income.group ) ) +
      geom_col() +
      facet_wrap( ~ year, ncol = 2 ) +
      labs( x = "Number of AI Conference Papers Per Capita",
            y = "Country",
            title = "Number of AI Conference Papers Per Capita by Country by Year" )
```



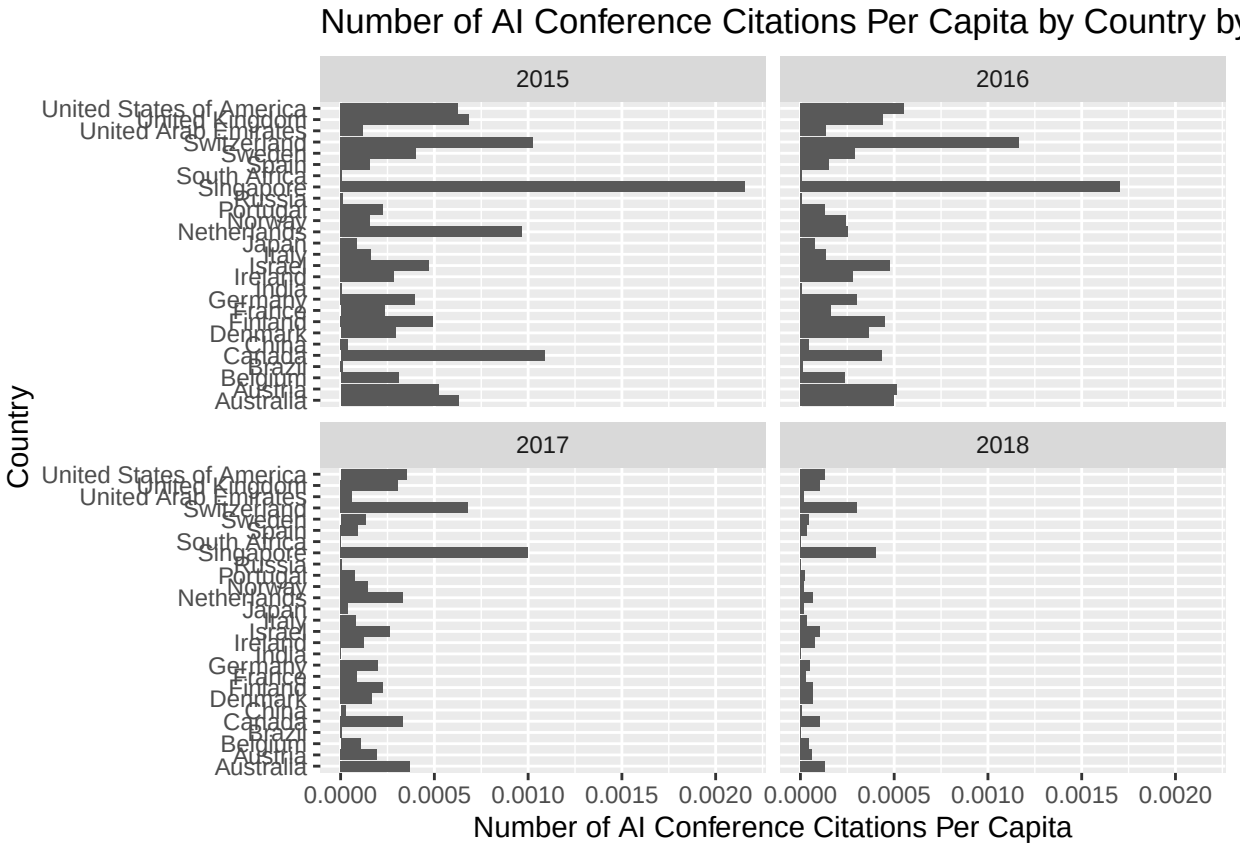

```
num_AIconf_citation_graph <- aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_citation, y = country ) ) +
      geom_col() +
      labs( x = "Number of AI Conference Citations",
            y = "Country",
            title = "Number of AI Conference Citations by Country" )


num_AIconf_citation_pc_graph <- aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_citation_pc, y = country ) ) +
      geom_col() +
      labs( x = "Number of AI Conference Citations Per Capita",
            y = "Country",
            title = "Number of AI Conference Citations Per Capita by Country" )

grid.arrange(num_AIconf_citation_graph, num_AIconf_citation_pc_graph, ncol = 2)
```




```
aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_citation, y = country ) ) +
      geom_col() +
      facet_wrap( ~ year, ncol = 2 ) +
      labs( x = "Number of AI Conference Citations",
            y = "Country",
            title = "Number of AI Conference Citations by Country by Year" )
```




```
aiVibrancyIndicators2019.tib %>%
  ggplot( aes( x = num_AIconf_citation_pc, y = country ) ) +
      geom_col() +
      facet_wrap( ~ year, ncol = 2 ) +
      labs( x = "Number of AI Conference Citations Per Capita",
            y = "Country",
            title = "Number of AI Conference Citations Per Capita by Country by Year" )
```


Number of AI Conference Citations Per Capita by Country by
2015
2016
2017
2018
Country
United States of America
United Kingdom
United Arab Emirates
Switzerland
Sweden
Spain
South Africa
Singapore
Russia
Portugal
Norway
Netherlands
Japan
Italy
Israel
Ireland
India
Germany
France
Finland
Denmark
China
Canada
Brazil
Belgium
Austria
Australia
0.0000
0.0005
0.0010
0.0015
0.0020
Number of AI Conference Citations Per Capita