|  |
| --- |
| Data Science Competency Training  Core Level Topic AssignmenT  EDA and visualisation |

Assignment Submission

Name: Tan Min Sheng Vincent

1. What are the trends that can be witnessed in the domestic and foreign market size indices?

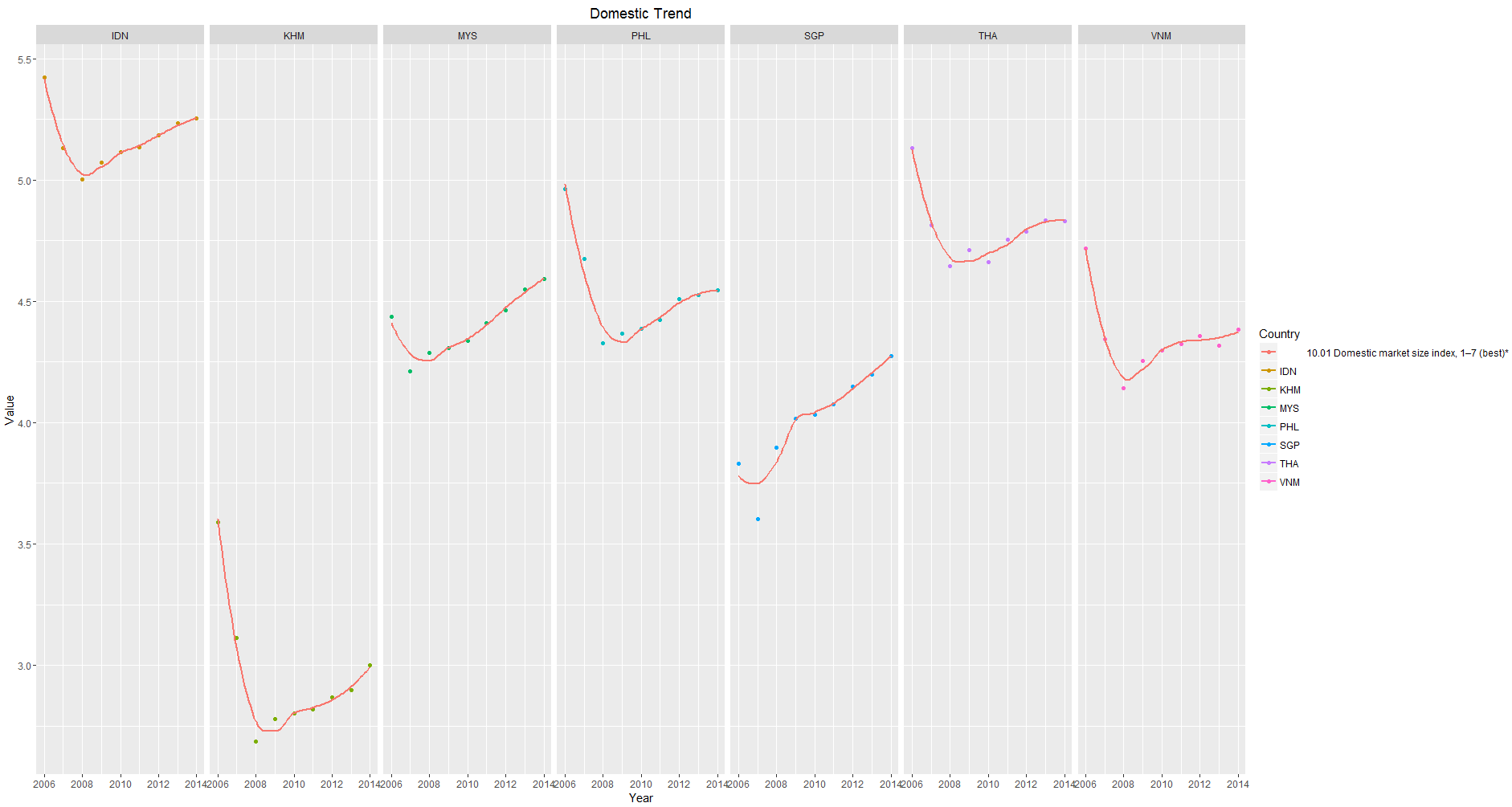
Columns Used:

**10.01** Domestic market size index   
Sum of gross domestic product plus value of imports of goods and services, minus value of exports of goods and services, normalized on a 1–7 (best) scale   
The size of the domestic market is calculated as the natural log of the sum of the gross domestic product valued at PPP plus the total value (PPP estimates) of imports of goods and services, minus the total value (PPP estimates) of exports of goods and services. Data are then normalized on a 1–7 scale.   
PPP estimates of imports and exports are obtained by taking the product of exports as a percentage of GDP and GDP valued at PPP.

**10.02** Foreign market size index

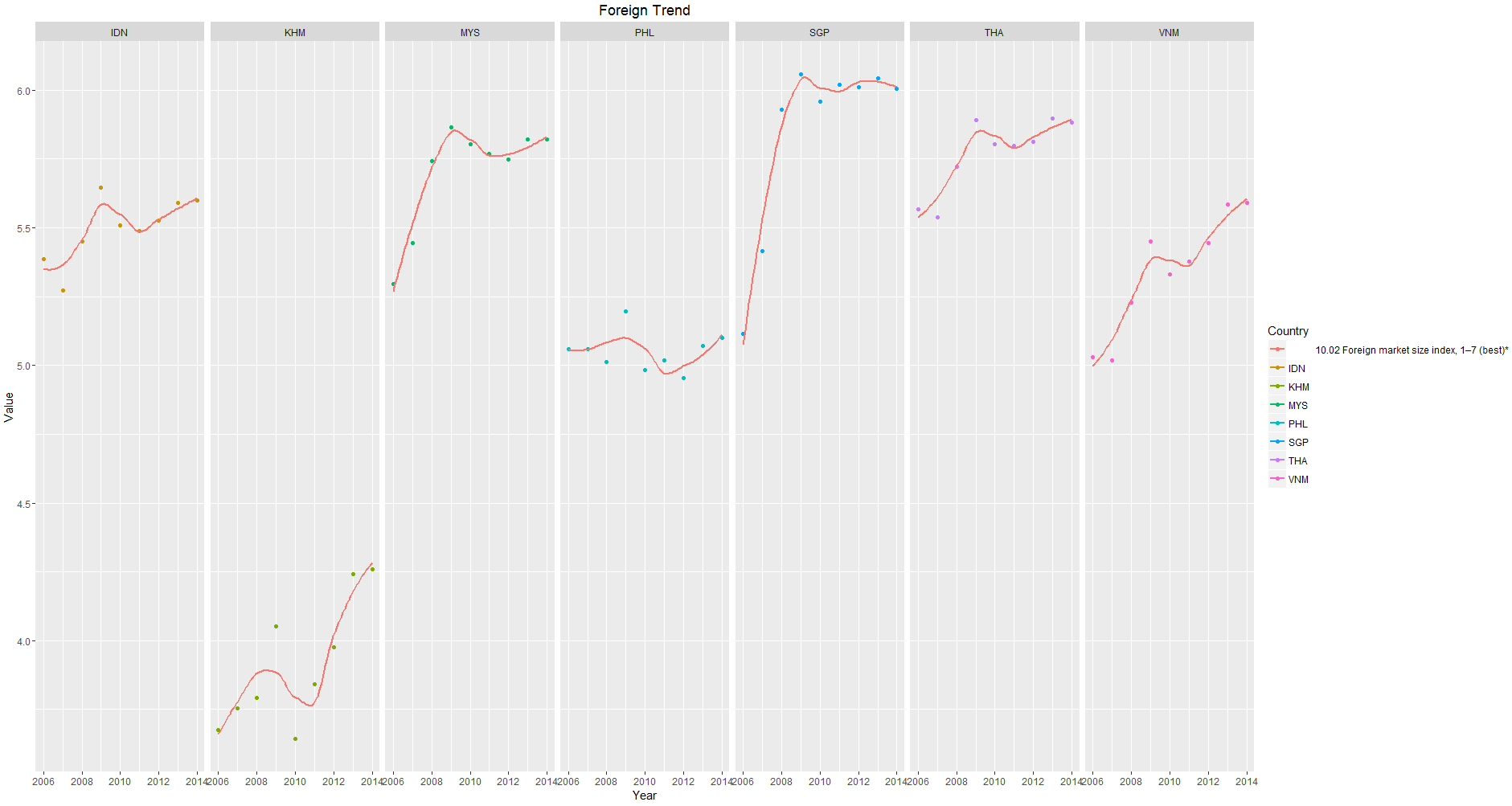
Value of exports of goods and services, normalized on a 1–7 (best) scale   
The size of the foreign market is estimated as the natural log of the total value (PPP estimates) of exports of goods and services, normalized on a 1–7 scale.   
PPP estimates of exports are obtained by taking the product of exports as a percentage of GDP and GDP valued at PPP.

Observations:

Domestic: 

* There was a significant decrease in the Domestic Market size across all the countries during 2006 – 2009.
* Followed by a slight increase in Domestic Market performance from 2009 – 2015
* Overall, the Domestic Market size from 2006 – 2015 had a slight down trend across most of the countries except Singapore and Malaysia.
* Singapore and Malaysia are the only countries that witnessed an overall uptrend.

Foreign:



* Overall, the Foreign Market size from 2006 – 2015 witnessed an uptrend across all countries.
* Singapore witnessed the most significant uptrend from 2006 – 2009, peaking at 2009.

Overall:



* The above graphs shows the domestic and foreign market size over the years.
* It can be observed that Cambodia lacks behind the rest of the countries in both markets sizes.



* The above graph shows a rough overview of both markets, domestic and foreign, over the years.
* Overall, most countries enjoyed an uptrend with the exception of Philippines, due to its decline in the foreign market.
* Singapore witnessed the sharpest climb, attributed by its strong foreign market.
* Indonesia bested Thailand into the spot of highest overall market score size over the years.

1. How does quality of the education system affect the GDP?

Columns Used:

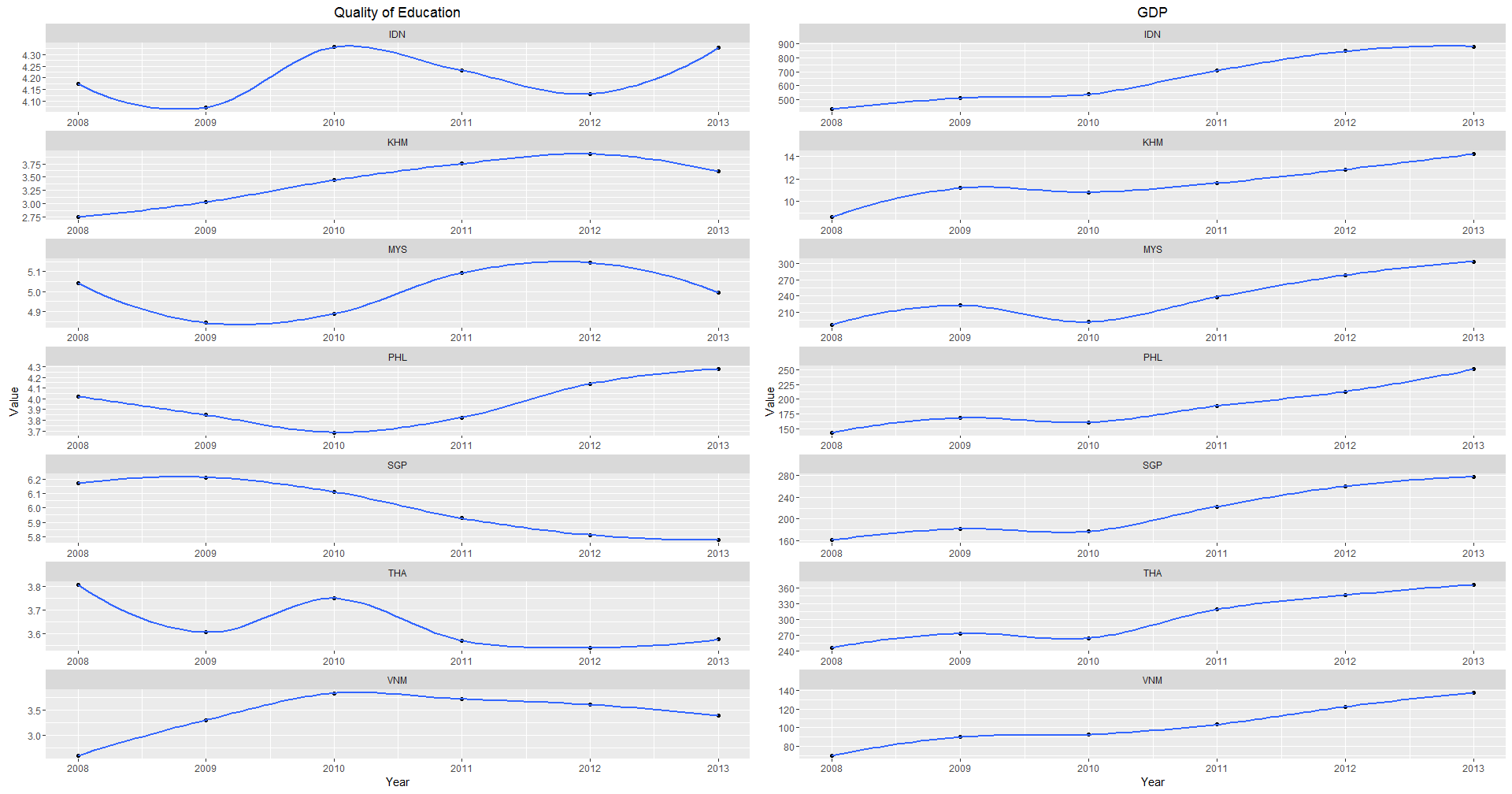
5.03 Quality of the education system   
How well does the education system in your country meet the needs of a competitive economy? [1 = not well at all; 7 = extremely well] weighted average

0.01 Gross domestic product  
Gross domestic product in billions of current US dollars

0.03 GDP per capita   
Gross domestic product per capita in current US dollars

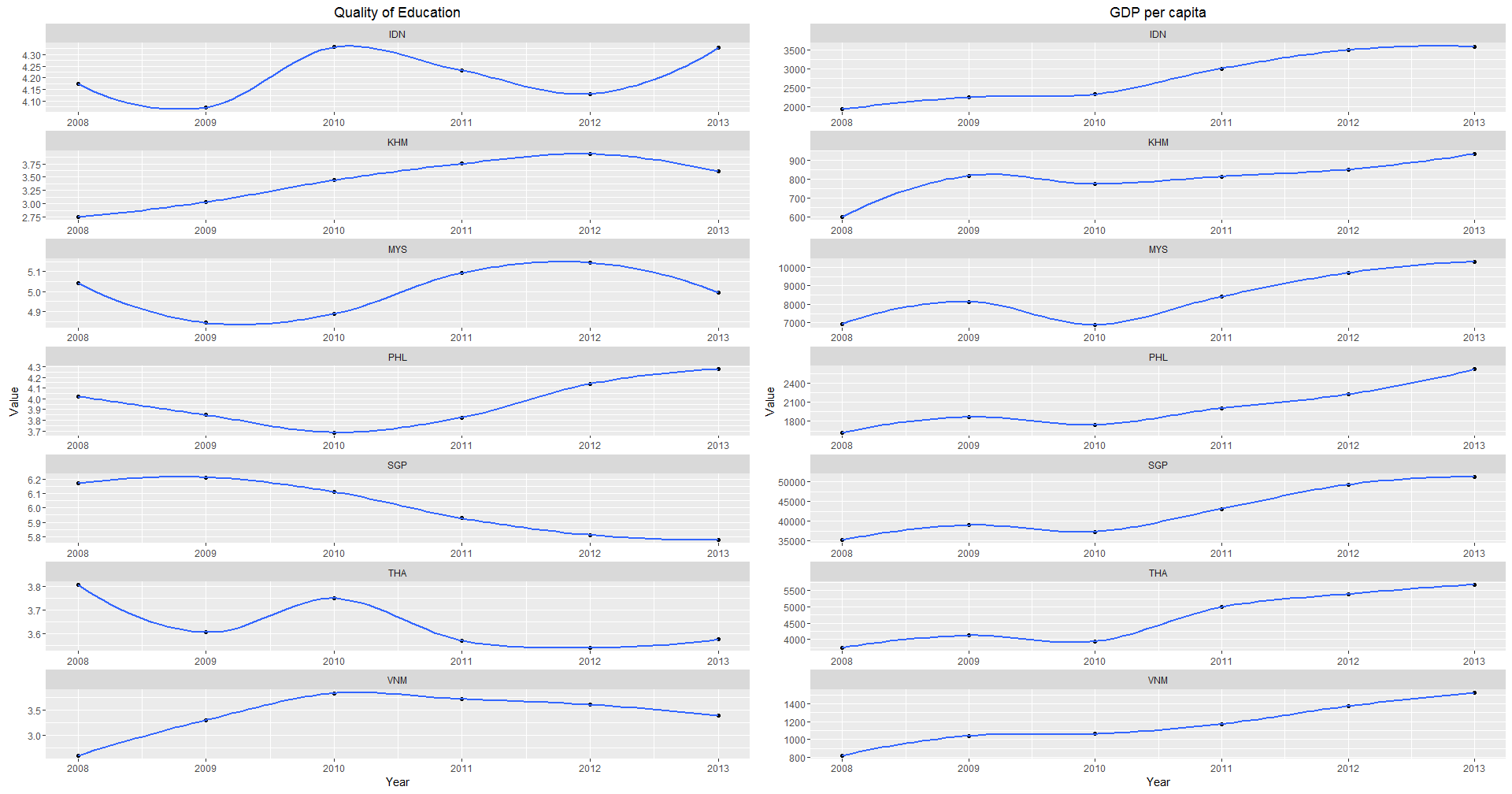
Observations:

#Quality of Education vs GDP



* The result of Pearson’s product-moment Correlation yield a correlation score of 0.1724164, showing low correlation between Quality of Education and GDP.
* The p-value of the Chi Square test yield result of 0.2379, which is greater than the .05 significance level, we do not reject the null hypothesis that the Quality of Education is independent of GDP.

#Quality of Education vs GDP per Capita



* The result of Pearson’s product-moment Correlation yield a correlation score of 0.8322916, showing strong correlation between Quality of Education and GDP per Capita.
* The p-value of the Chi Square test yield result of 0.4828 is greater than the .05 significance level, we do not reject the null hypothesis that the Quality of Education is independent of GDP.

#Overall

* Through the above two graphs, it can be observed that there is no direct causation from Quality of Education Systems to GDP, nor GDP per Capita, despite significantly high correlation.
* Which coincides with my Knowledge of the world. (aka Google)
  + <http://www.investopedia.com/articles/economics/09/education-training-advantages.asp>
  + <http://educationnext.org/education-and-economic-growth/>
* I conclude that Quality of Education Systems does not directly affect GDP. However I do not rule out that it might be one of the contributing factors affecting GDP.

1. Does lower level of education lead to more HIV prevalence for year 2006 and 2015?

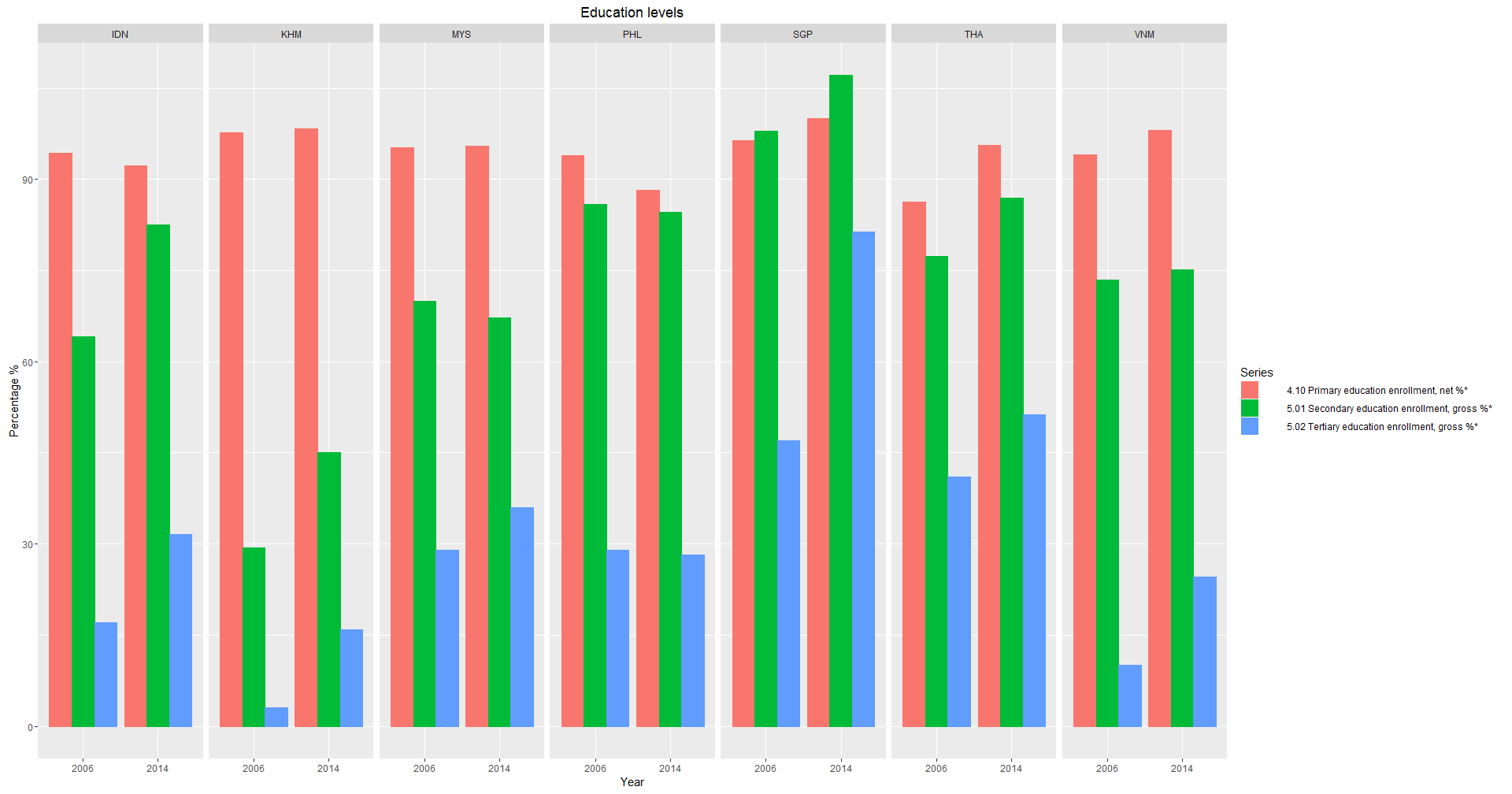
Columns Used:

5.01 Secondary education enrollment rate   
Gross secondary education enrollment rate  
Reported value corresponds to the ratio of total secondary enrollment, regardless of age, to the population of the age group that officially corresponds to the secondary education level.   
  
5.02 Tertiary education enrollment rate   
Gross tertiary education enrollment rate  
Reported value corresponds to the ratio of total tertiary enrollment, regardless of age, to the population of the age group that officially corresponds to the tertiary education level.

4.05 HIV prevalence   
HIV prevalence as a percentage of adults aged 15–49 years  
HIV prevalence refers to the percentage of people aged 15–49 who are infected with HIV at a particular point in time, no matter when infection occurred.

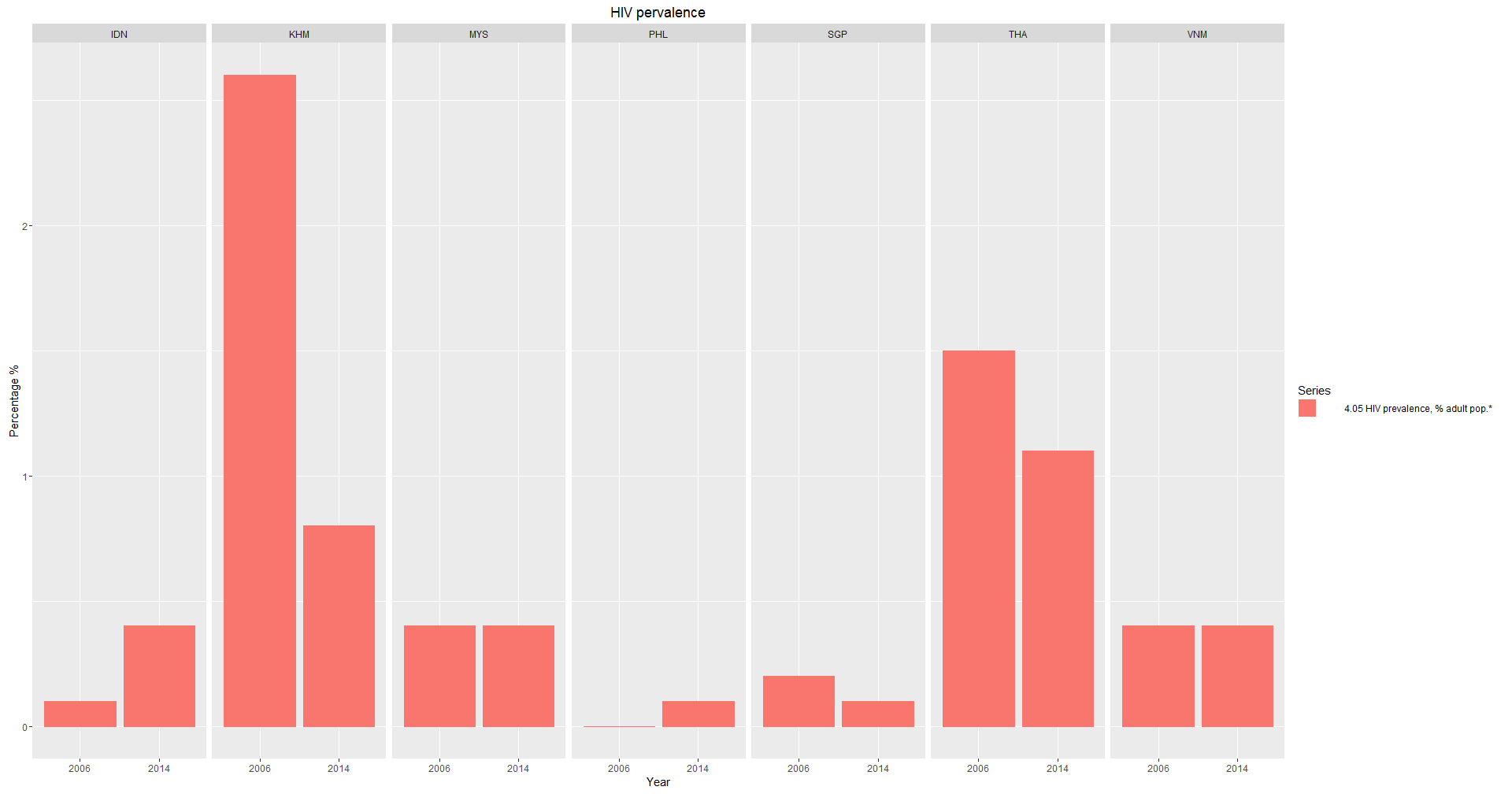
Observations:

#Education Level



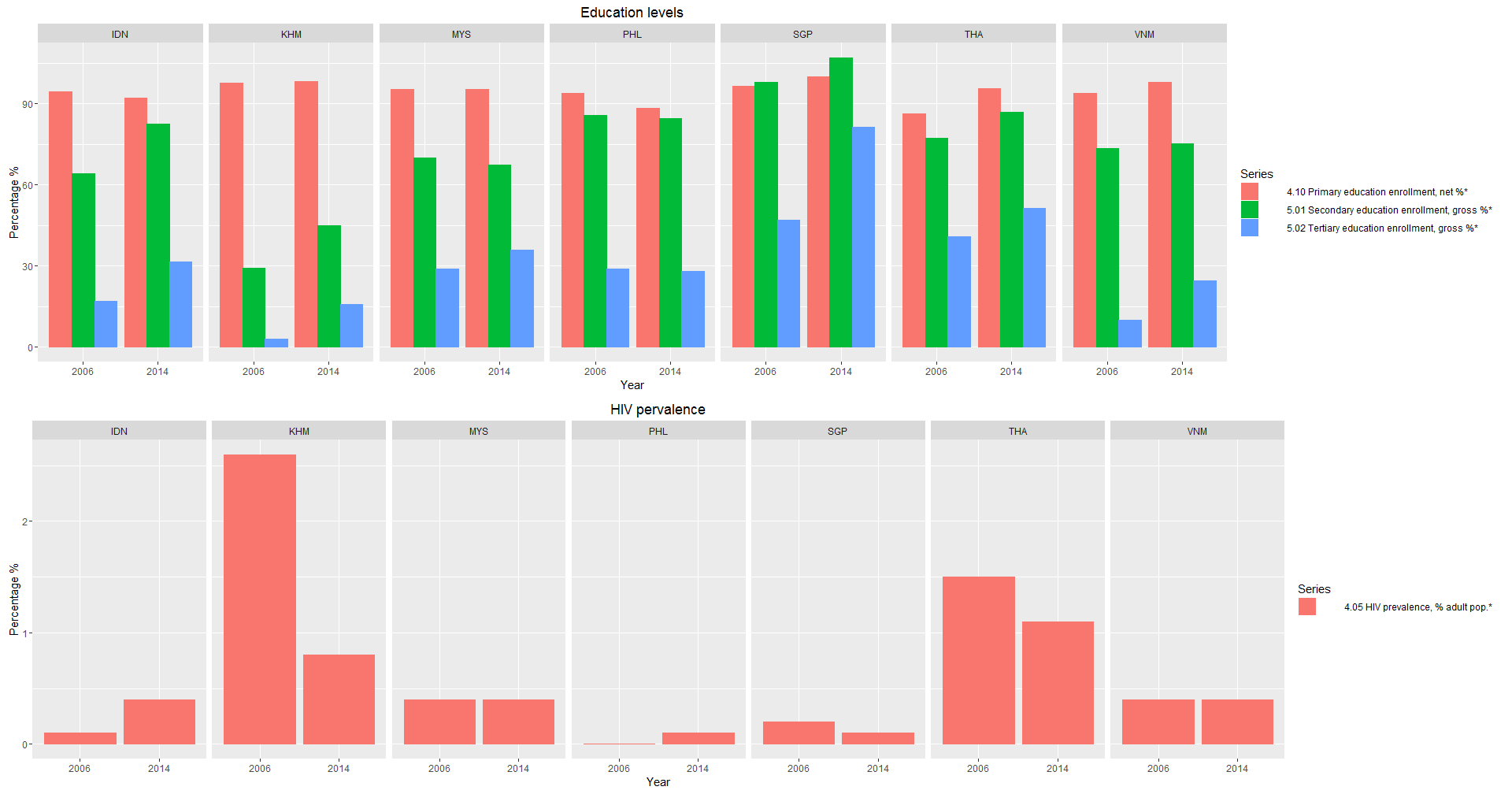
* The general downward trend in the percentage of student enrollment from primary to tertiary education can be witnessed across all the countries and years except Singapore. This is to be expected as most countries provide subsidies or basic education for their citizens, but with less higher level educational institutions, the cost of education increases with its level, less families can provide their children.
* For both 2006 and 2014, Singapore seen peaks in the secondary education enrollment. Which might be caused by the influx of foreign talents and their families.
* Most countries also witness an increase in enrollment for higher education in 2014 when compared to enrollment in 2006.

#HIV prevalence



* HIV prevalence have a downward trend from 2006 to 2014 for Cambodia, Singapore and Thailand.
* Malaysia and Vietnam witnessed the same level of HIV prevalence from 20016 to 2014.
* Indonesia and Philippines saw a noticeable increase in 2006 to 2014 in the same period.

#Education Level vs HIV Prevalence



* Countries that witnessed a decrease in HIV prevalence also witnessed a significant increase in secondary education enrollment from 2006 to 2014.
* Countries that witness stagnate HIV prevalence did not see as much change in their secondary education enrollment in the same period.
* The decrease in secondary education enrollment seen from Philippines coincides with its increase in HIV prevalence over the same period. However, the same could not be sad for Indonesia with increase in both enrollment rates and HIV prevalence from 2006 to 2014.
* I conclude that education level is not the primary factor, but might be one of many the factors causing HIV prevalence.

1. Does irregular payments and bribes as well as diversion of public funds lead to lower / higher reliability of police services for the last 5 years?

Columns Used:

1.03 Diversion of public funds

In your country, how common is diversion of public funds to companies, individuals, or groups due to corruption? [1 = very commonly occurs; 7 = never occurs]

1.05 Irregular payments and bribes

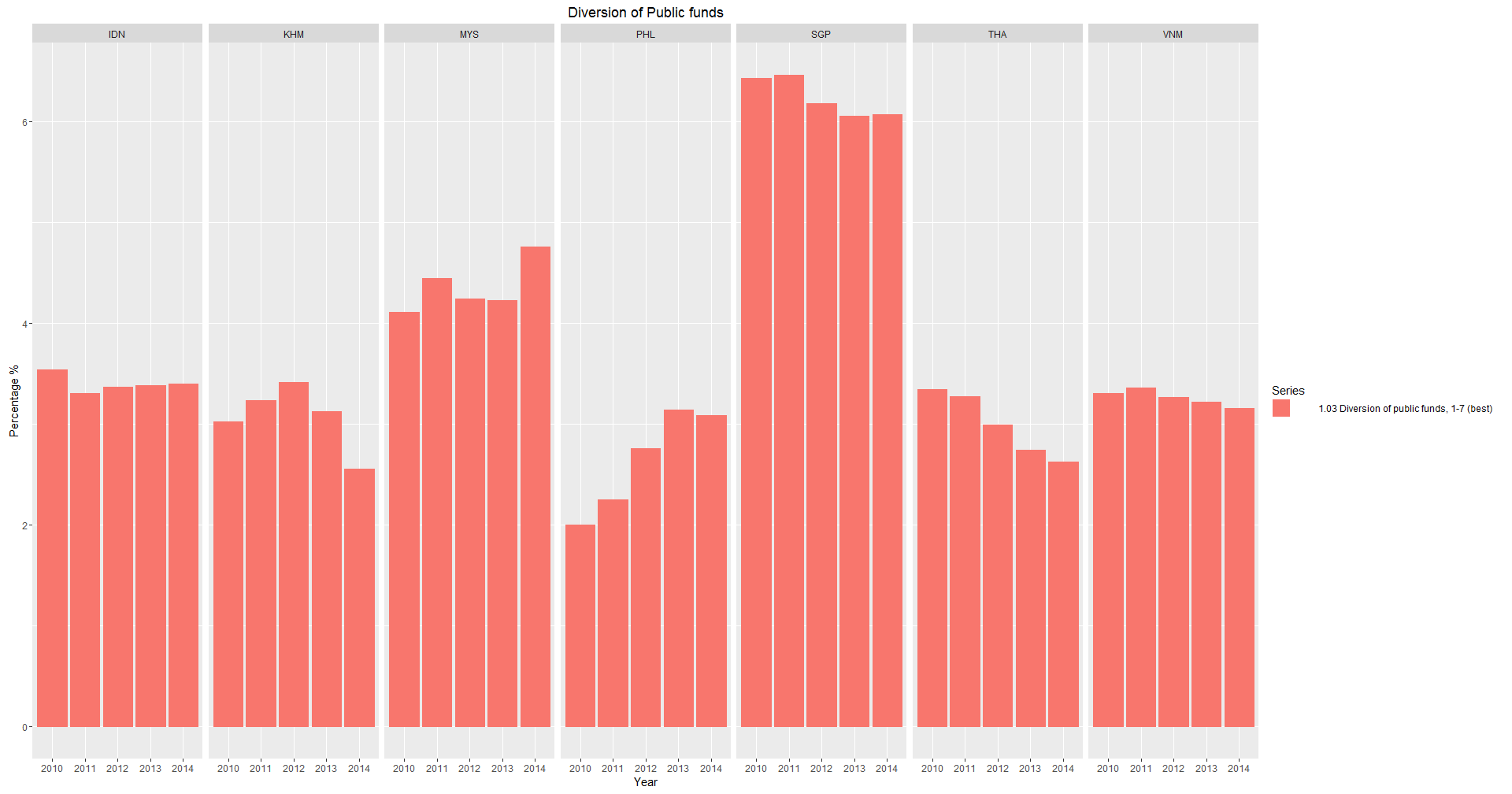
Average score across the five components of the following Executive Opinion Survey question: In your country, how common is it for firms to make undocumented extra payments or bribes in connection with (a) imports and exports; (b) public utilities; (c) annual tax payments; (d) awarding of public contracts and licenses; (e) obtaining favorable judicial decisions? In each case, the answer ranges from 1 [very common] to 7 [never occurs].

1.16 Reliability of police services

In your country, to what extent can police services be relied upon to enforce law and order? [1 = cannot be relied upon at all; 7 = can be completely relied upon]

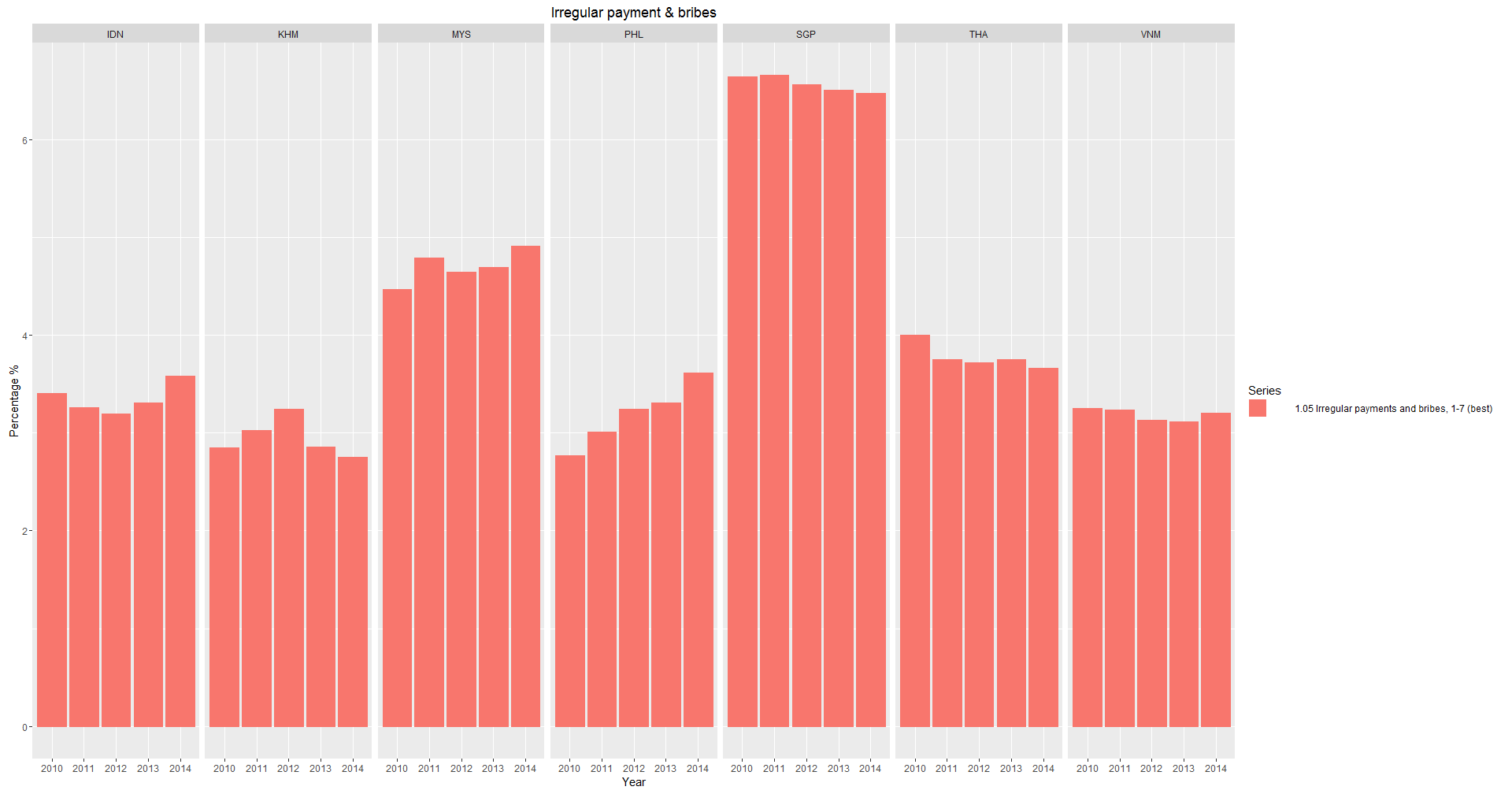
Observations:

# Diversion of Public Funds



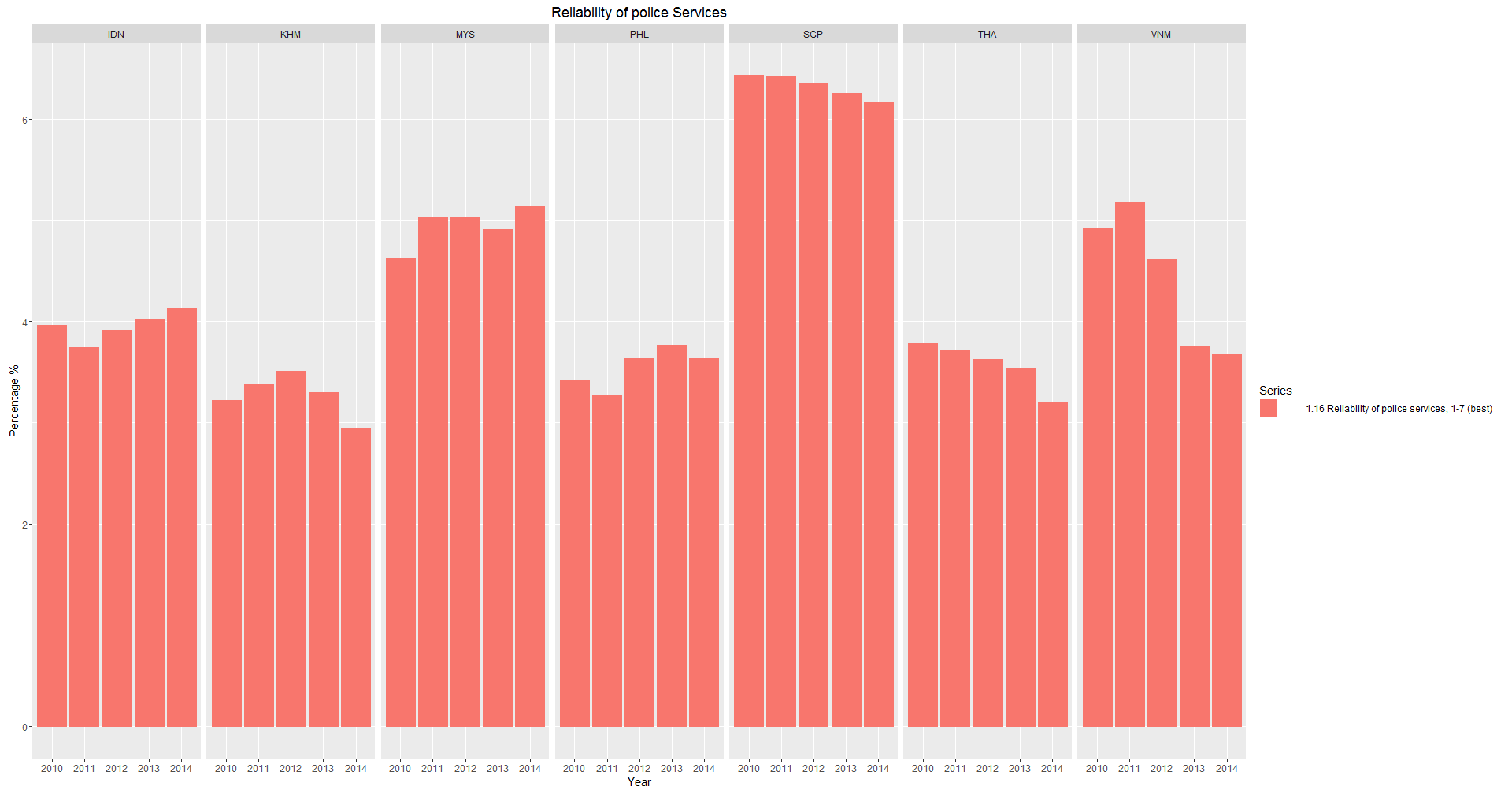
* It can be observed that most countries are rated around 2 – 4 with the exception of Malaysia (4-5) and Singapore (6-7).
* No obvious trends can be spotted between the countries.
* Philippines improved the most in the ratings over the years, peaking in 2013.

# Irregular payment & bribes



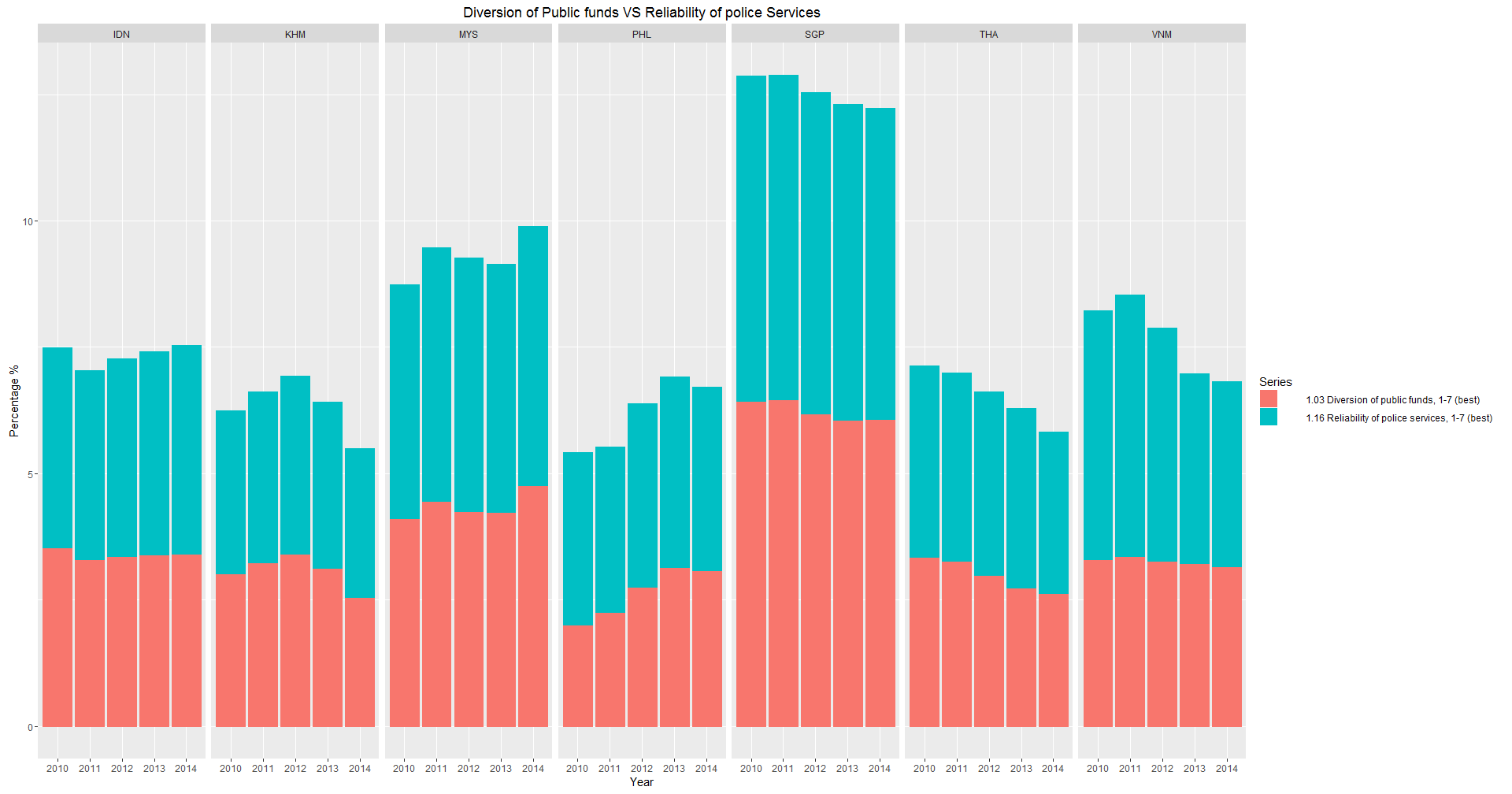
* It can be observed that most countries are rated around 2 – 4 with the exception of Malaysia (4-5) and Singapore (6-7).
* No obvious trends can be spotted between the countries.
* Philippines improved the most in the ratings over the years, peaking in 2014.

# Reliability of Police Services



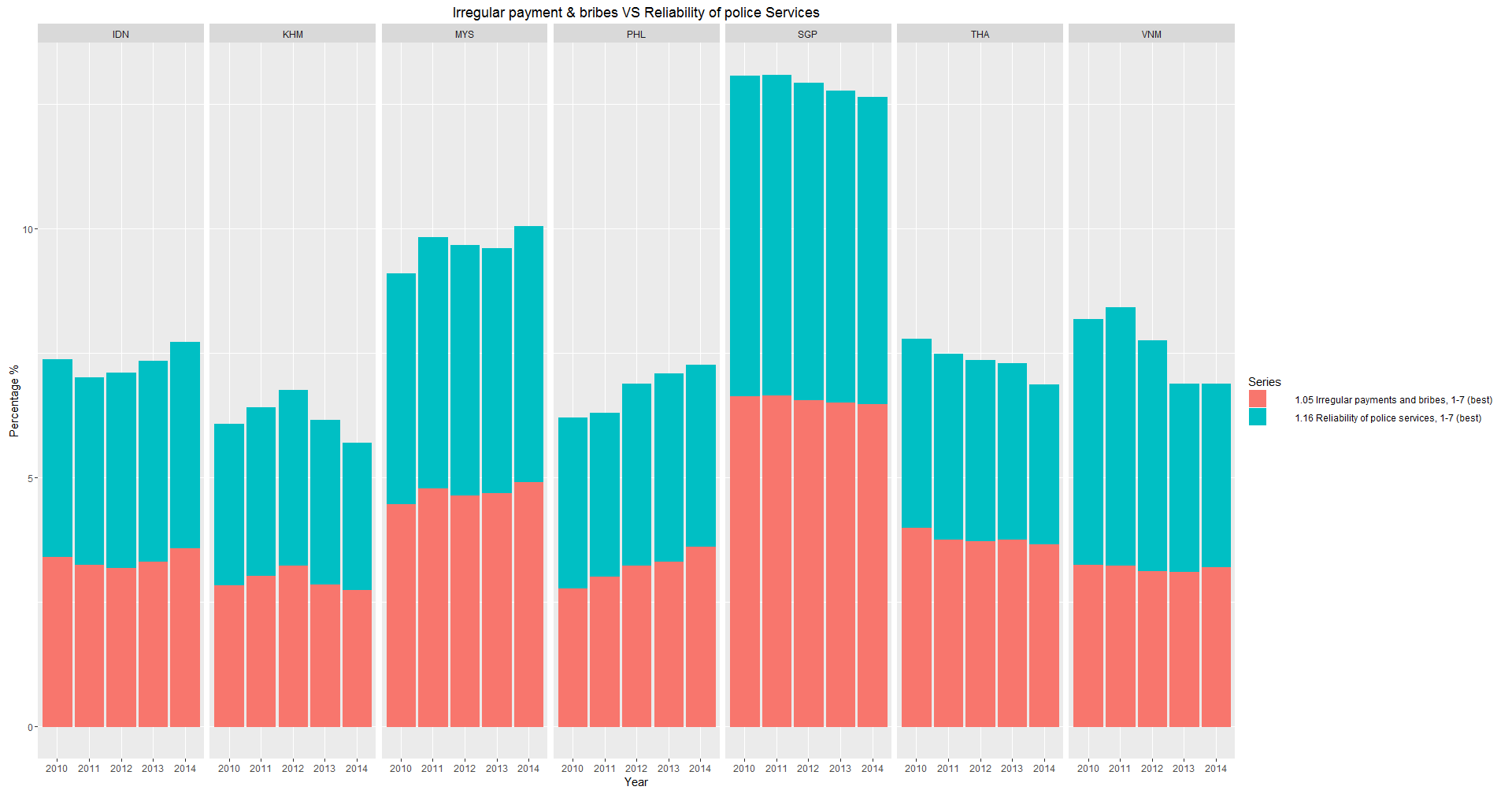
* It can be observed that Cambodia, Philippines and Thailand are rated closely.
* Singapore towers over the rest of the countries with a minimum of 1 point lead on the closest country.
* Vietnam’s ratings deteriorated rapidly after 2011, only to slow down from 2013 to 2014.

# Diversion of Public Funds vs Reliability of Police Services



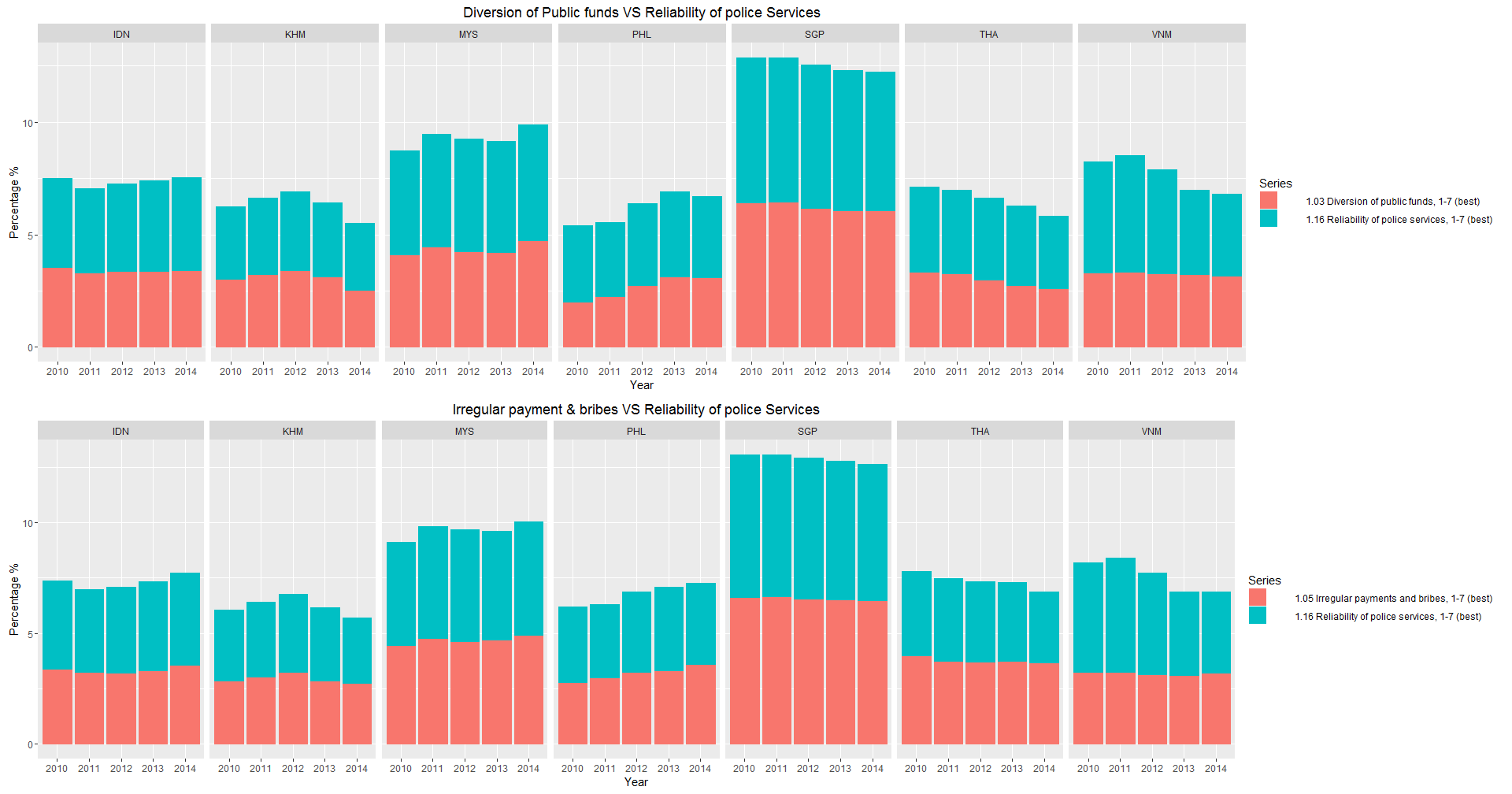
* The result of Pearson’s product-moment Correlation yield a correlation score of 0.9320889, showing strong correlation between Diversion of Public Funds and Reliability of Police Services.
* The p-value of the Chi Square test yield result of 0.2376 is greater than the .05 significance level, we do not reject the null hypothesis that the Diversion of Public Funds is independent of Reliability of Police Services.

# Irregular payment & bribes vs Reliability of Police Services



* The result of Pearson’s product-moment Correlation yield a correlation score of 0.9003074, showing strong correlation between Irregular Payment & Bribes and Reliability of Police Services.
* The p-value of the Chi Square test yield result of 0.2376 is greater than the .05 significance level, we do not reject the null hypothesis that the Irregular Payment & Bribes is independent of Reliability of Police Services.

# Overall



* From the findings above, we can conclude that Irregular Payment & Bribes, Diversion of Public Funds and Reliability of Police Services are highly correlated but are independent from each another. Thus do not lead to lower / higher reliability of police services for the last 5 years.

1. How has inflation and the soundness of banks affected the average income per person for each country?

Columns Used:

0.03 GDP per capita   
Gross domestic product per capita in current US dollars

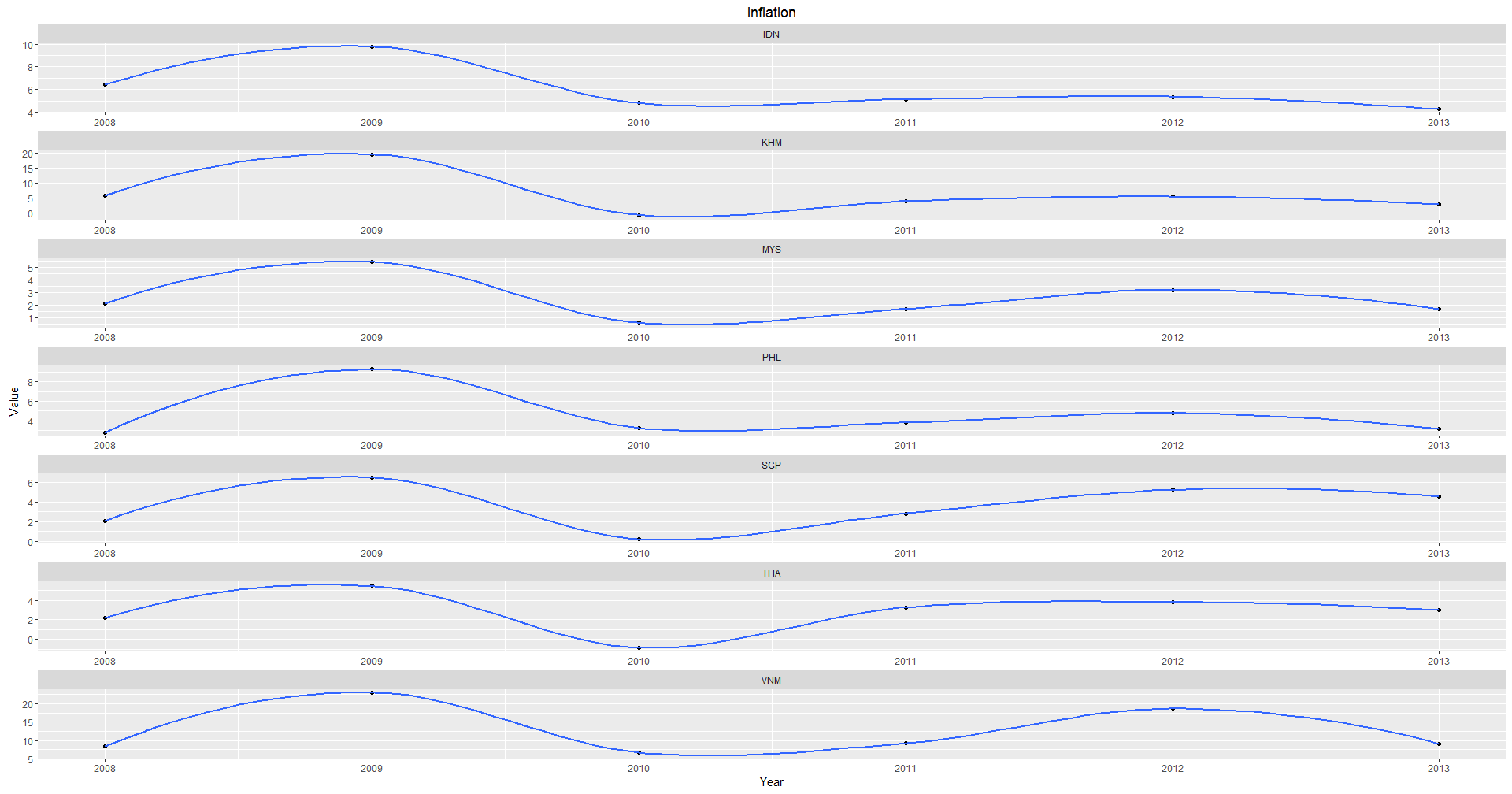
3.03 Inflation   
Annual percent change in consumer price index (year average)

For inflation rates between 0.5 and 2.9 percent, a country receives the highest possible score of 7. Outside this range, scores decrease linearly as they move away from these values.

8.06 Soundness of banks   
In your country, how would you assess the soundness of banks? [1 = extremely low—banks may require recapitalization; 7 = extremely high—banks are generally healthy with sound balance sheets]

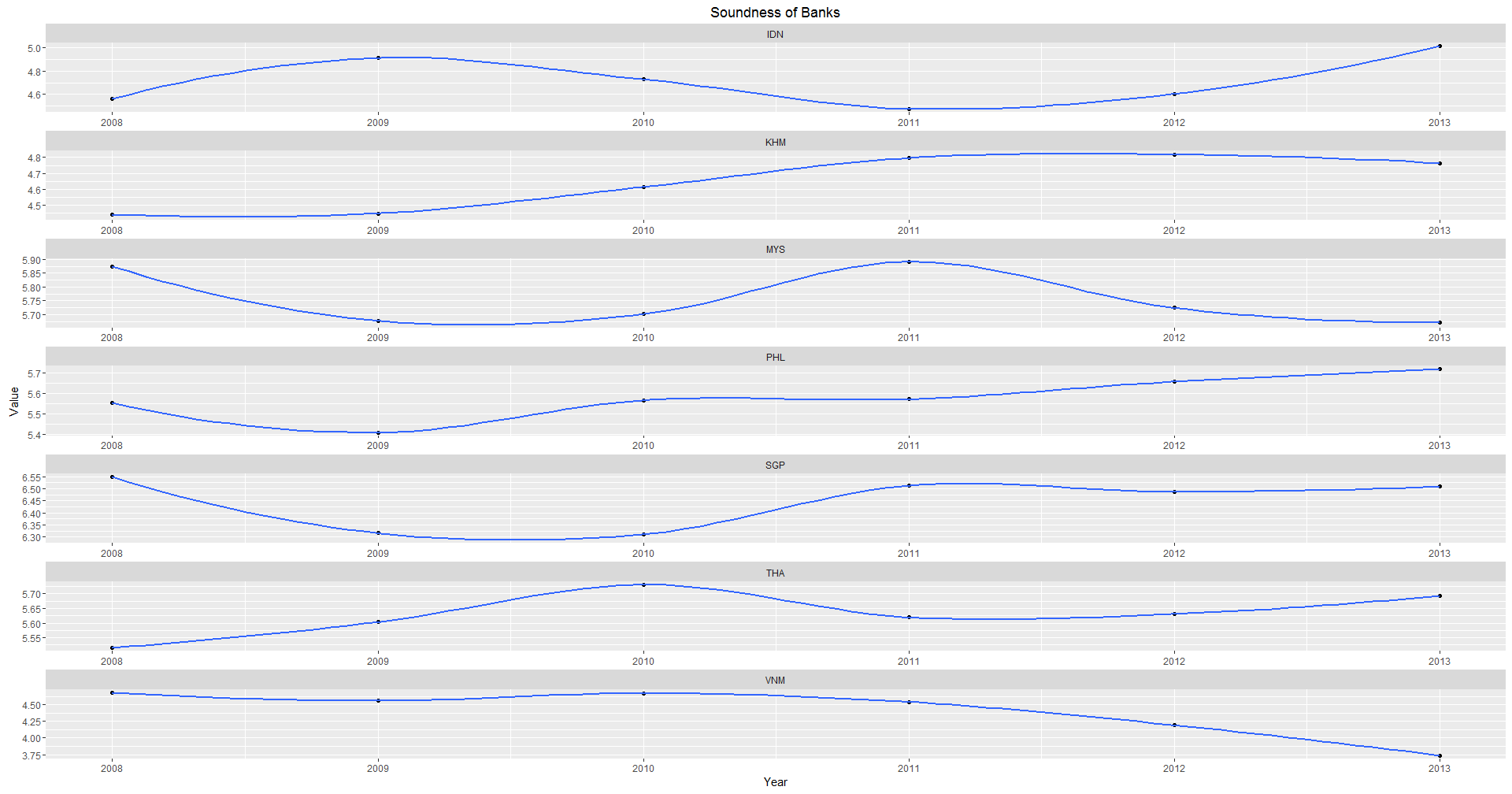
Observations:

# Inflation



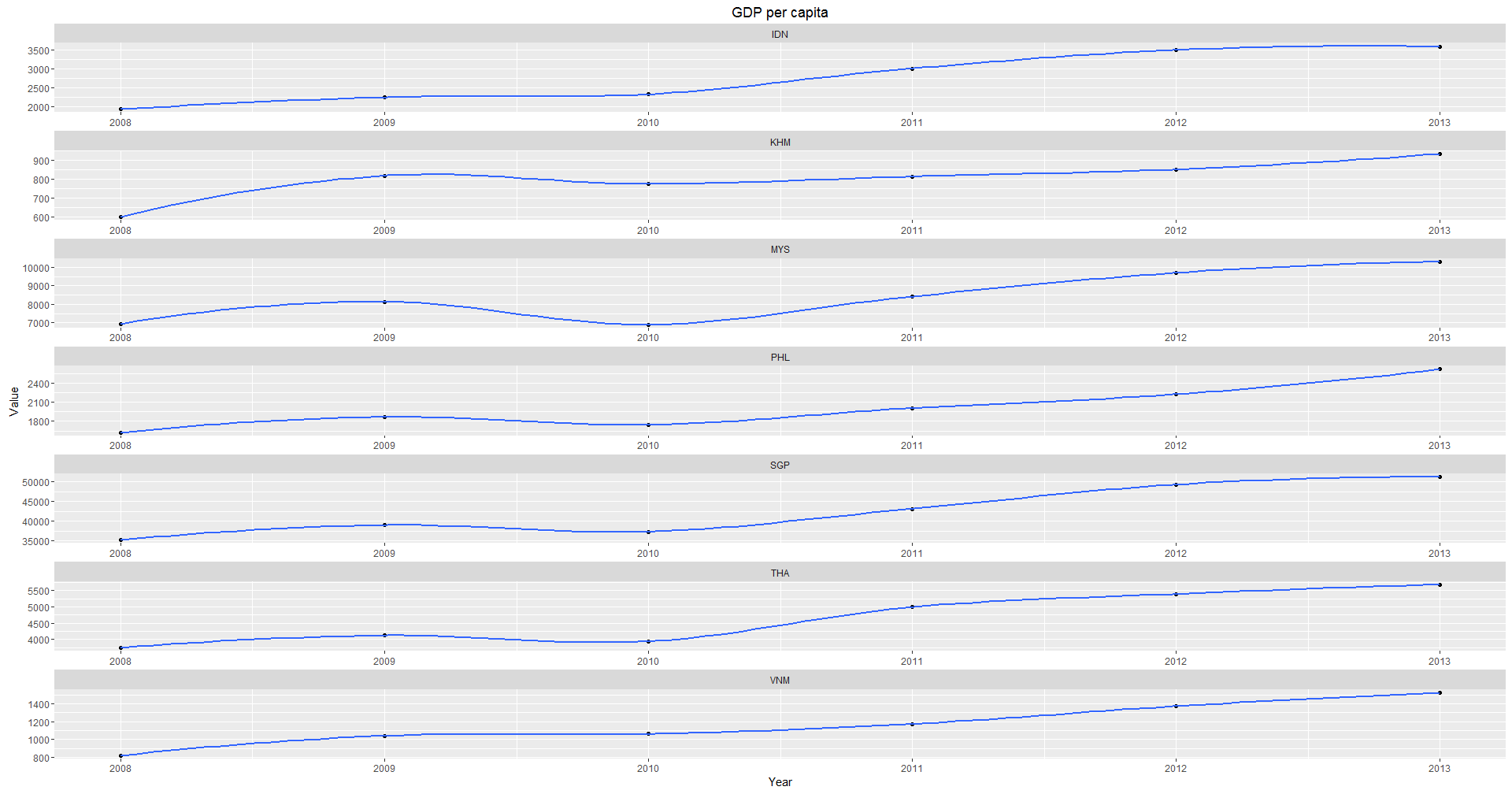
* The general down trend of inflation across all the countries resembles the letter ‘M’, with peaks seen in 2009 and 2012.

# Soundness of Banks



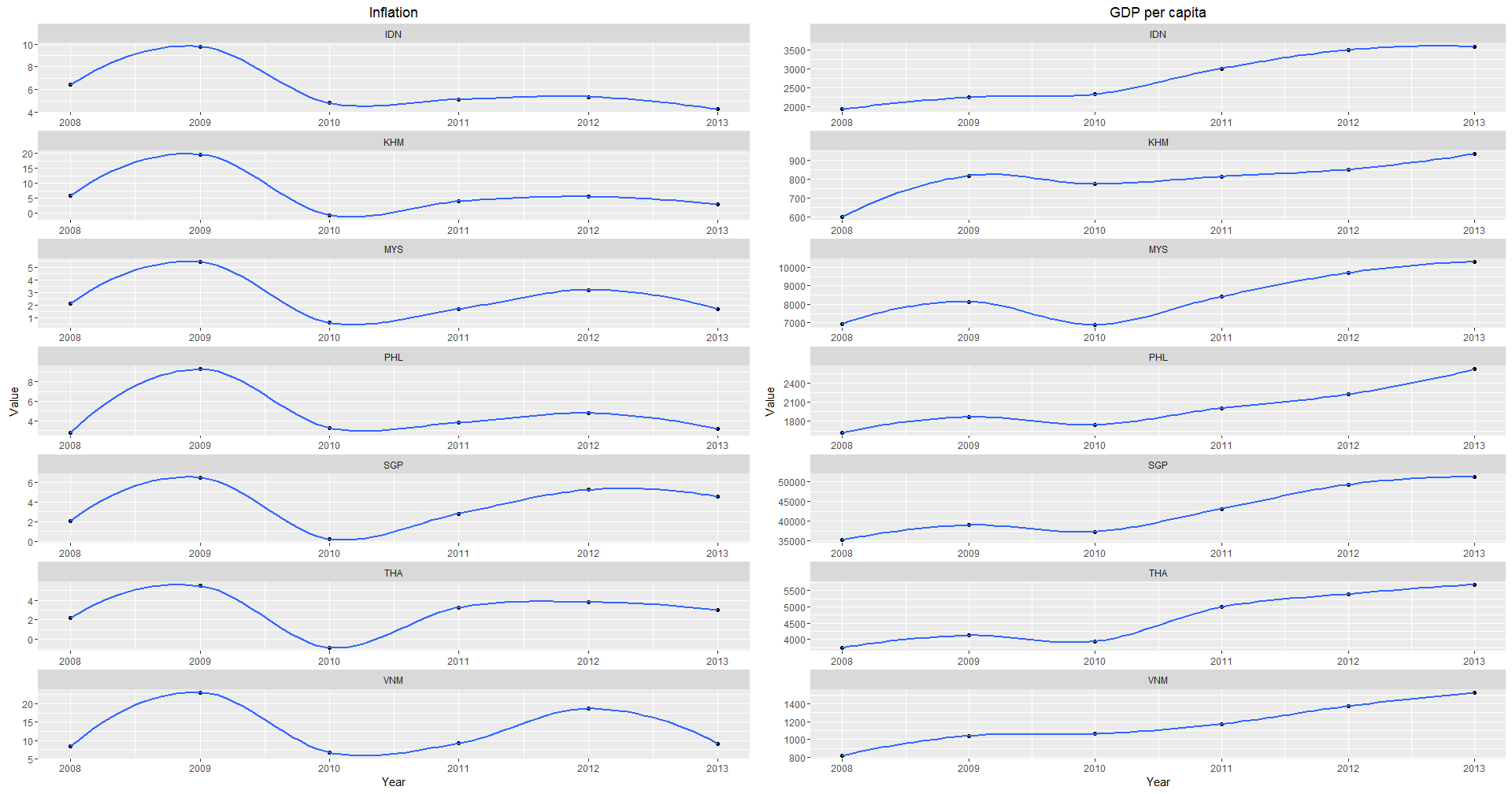
* No significant trend observed across the countries.

# GDP per Capita



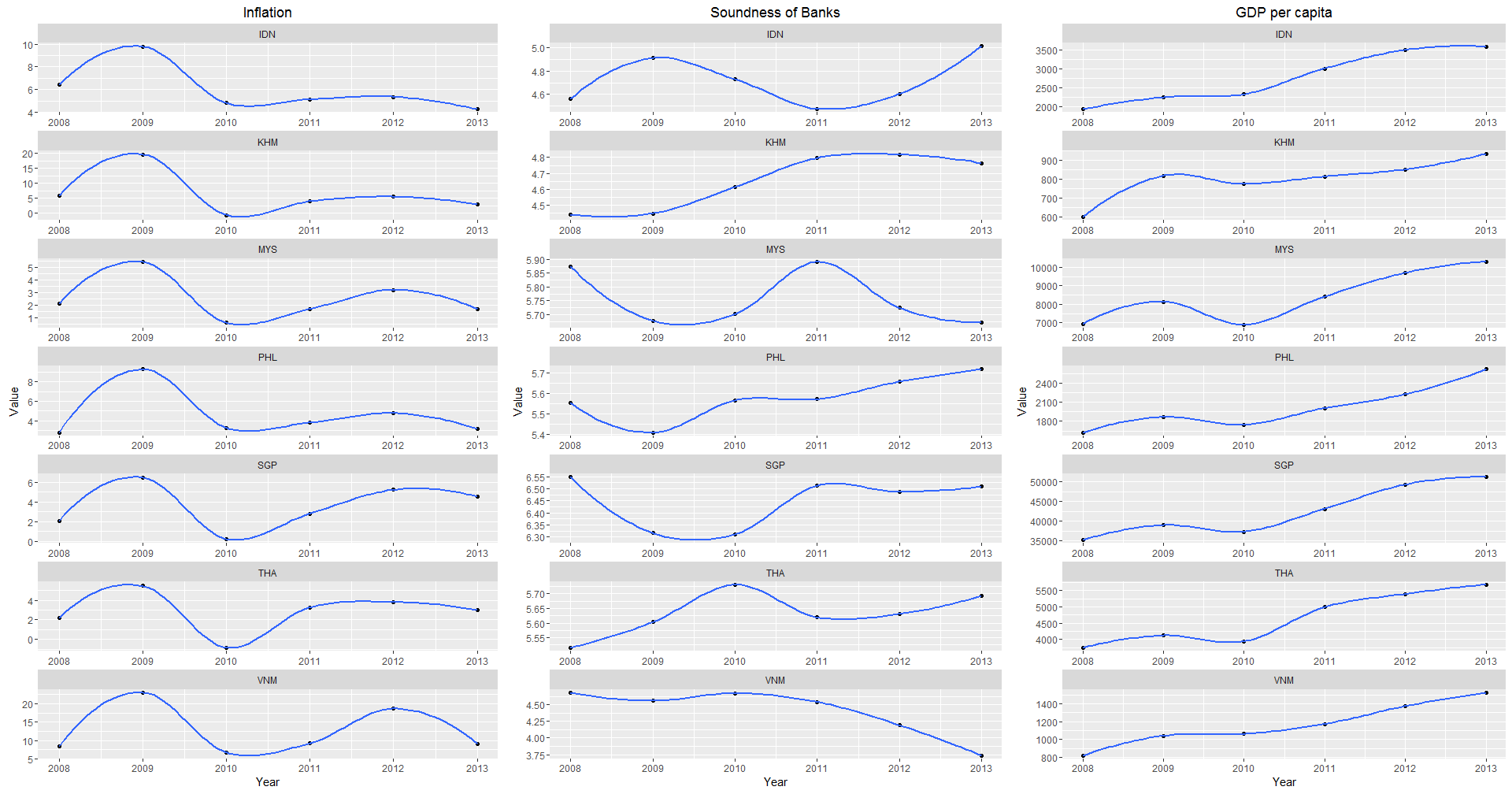
* A general upward trend can be seen for GDP per capita across all the countries.
* A significant disparity between GDP per capita of Singapore and the other countries can be observed.

# Inflation vs GDP per Capita



* GDP per Capita mimics Inflation across all countries from 2008 to 2010 with both inflation and GDP per Capita peaking in 2009 before it decreased again in 2010.
* The result of Pearson’s product-moment Correlation yield a correlation score of -0.2090328, showing weak negative correlation between Inflation and GDP per Capita.
* The p-value of the Chi Square test yield result of 0.2379 is greater than the .05 significance level, we do not reject the null hypothesis that the Inflation is independent of GDP per Capita.

# Inflation vs GDP per Capita



* The result of Pearson’s product-moment Correlation yield a correlation score of 0.7293571, showing positive correlation between Bank Soundness and GDP per Capita.
* The p-value of the Chi Square test yield result of 0.2379 is greater than the .05 significance level, we do not reject the null hypothesis that the Bank Soundness is independent of GDP per Capita.
* I conclude that inflation and the soundness of banks do not directly affect the average income per person for each country.
* However it does not rule out the possibility that they might be contributing factors to GDP per capita.