

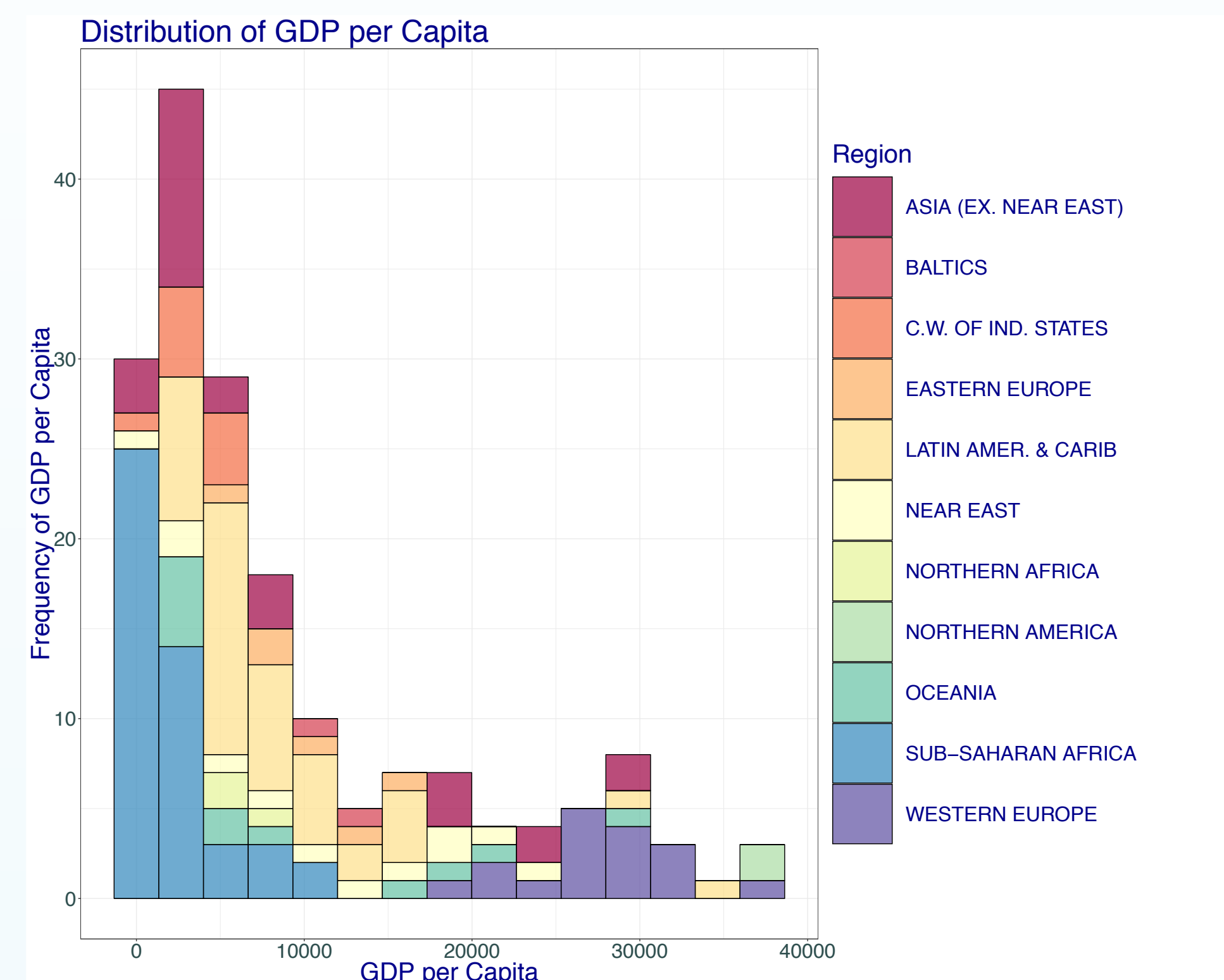
36-315 Project Group 2

“Countries of The World”

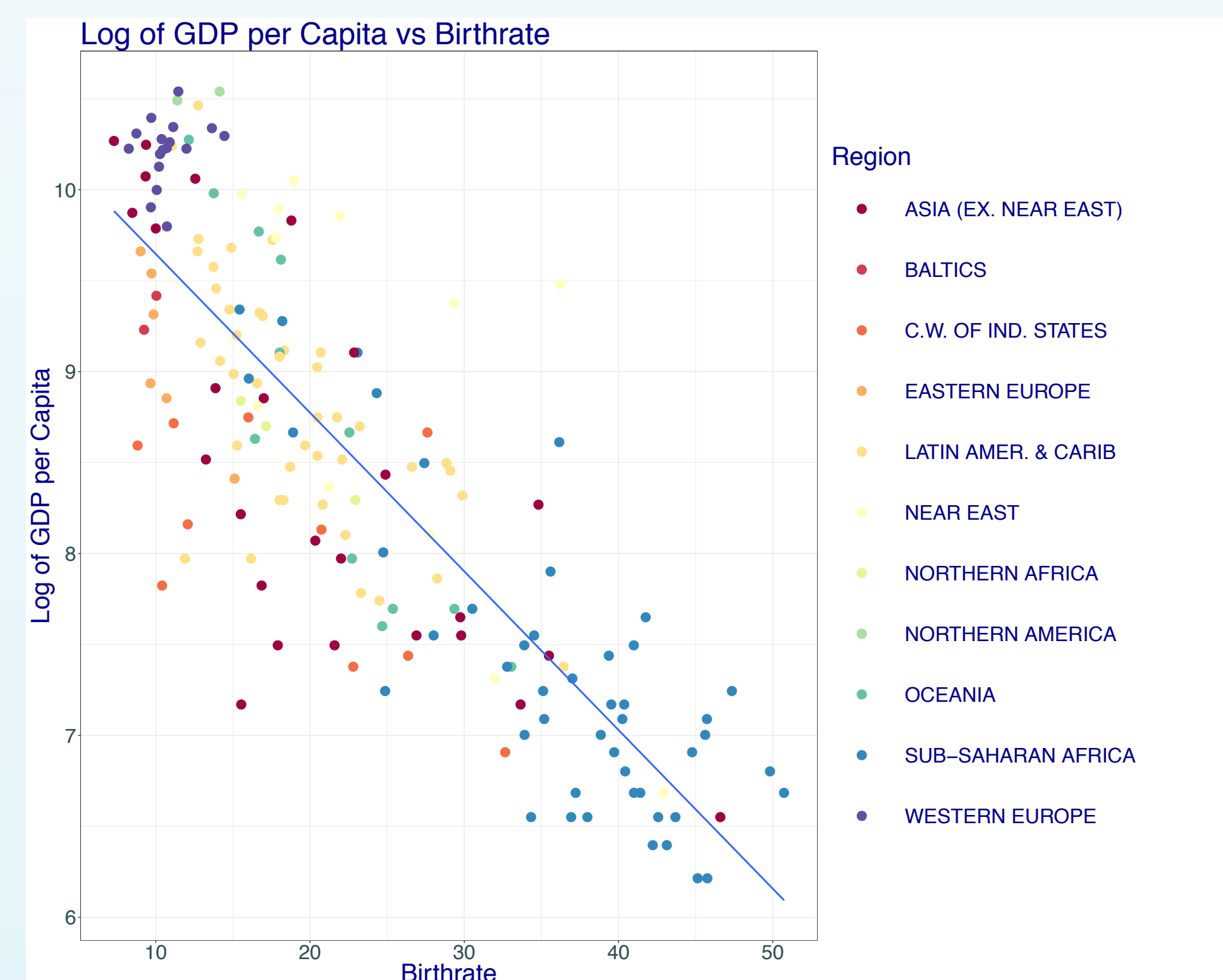
Wanhe Zhao, Minami Makino, Shreya Prakash, Masahero Masuda
Carnegie Mellon University Department of Statistics

Introduction

We decided to look at the Countries of the World Dataset on Kaggle. The dataset contains observations corresponding to specific countries of the world. For each country, we have GDP, economic and demographic information. We will explore different factors that contribute to a country's GDP. We hypothesized that the higher GDP countries would have higher literacy rate, lower birth rate and higher proportions of service sectors. It is also intuitive that GDP and those factors vary per country and region.

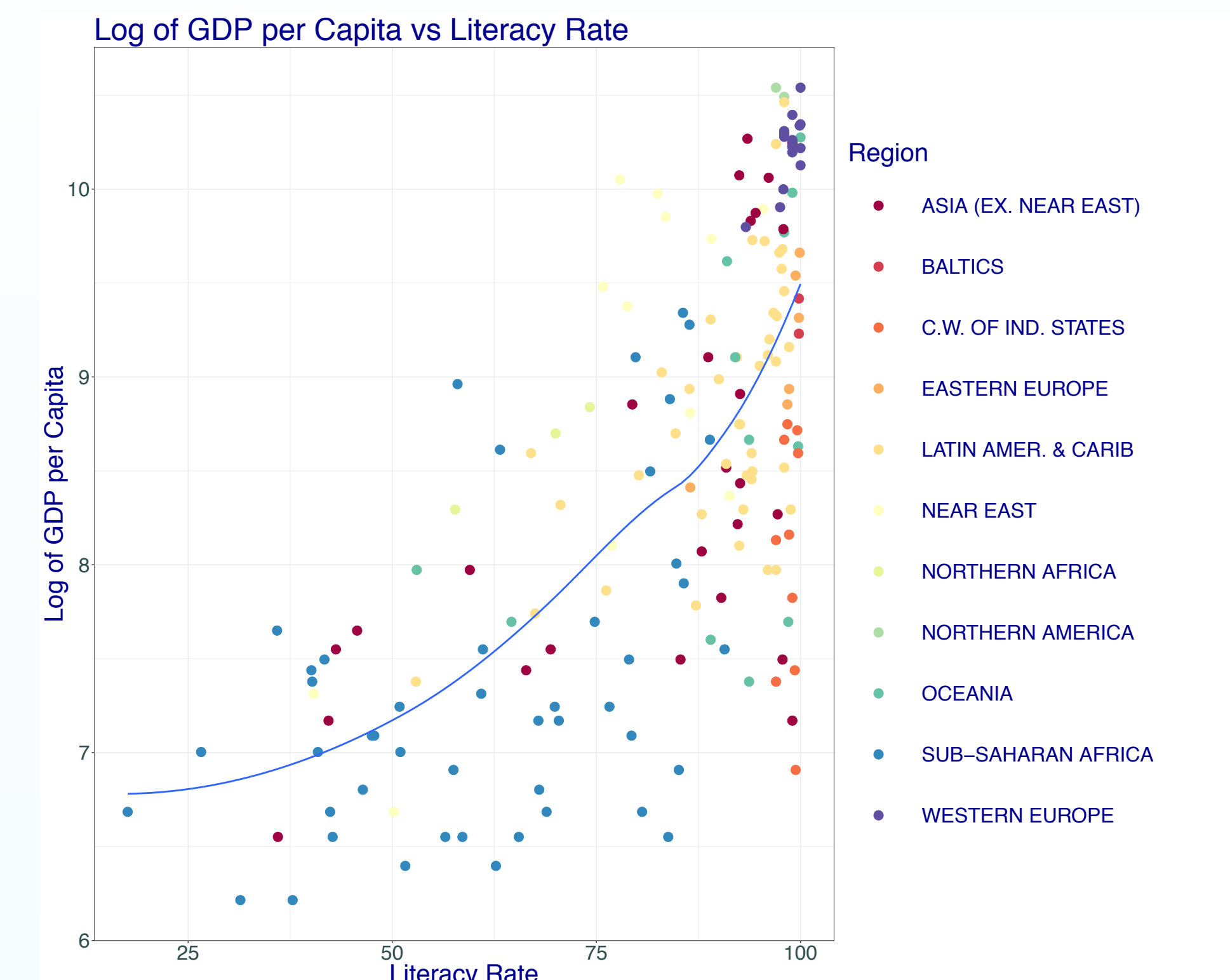


- The distribution of GDP per Capita appears to be skewed to the right, with majority of regions having a GDP per Capita of <10000.
- It is important to note that Sub-Saharan Africa and Latin America consist of the majority of the lower GDP per Capita, with the higher ends encompassing Northern America and Western Europe.

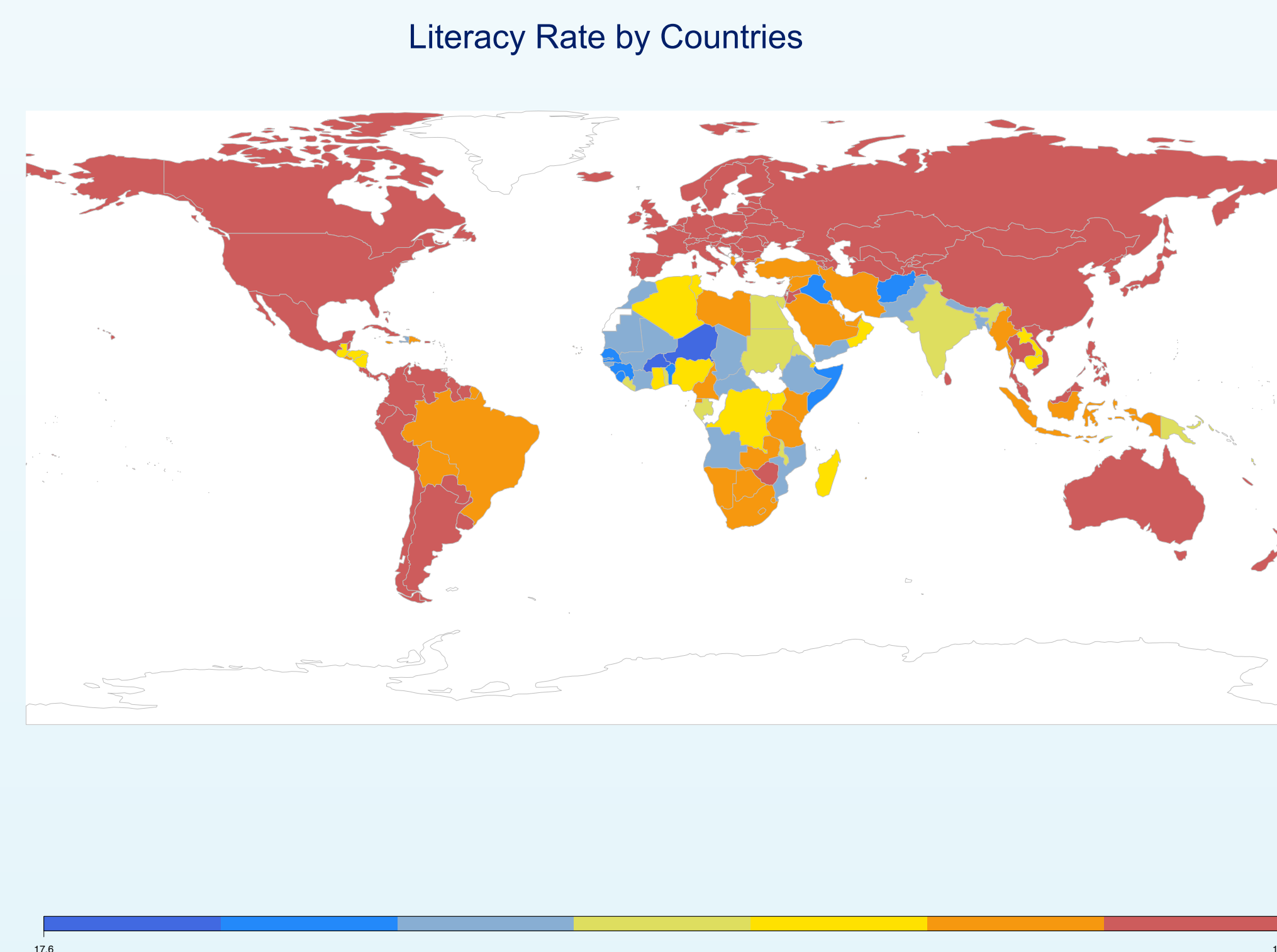


- There appears to be a negative linear correlation between birth rate and GDP per capita.
- Regions like Sub-Saharan Africa have high birth rate and low log GDP per capita.
- Regions like Western Europe have low birth rate and high log GDP per capita.

Literacy Rate

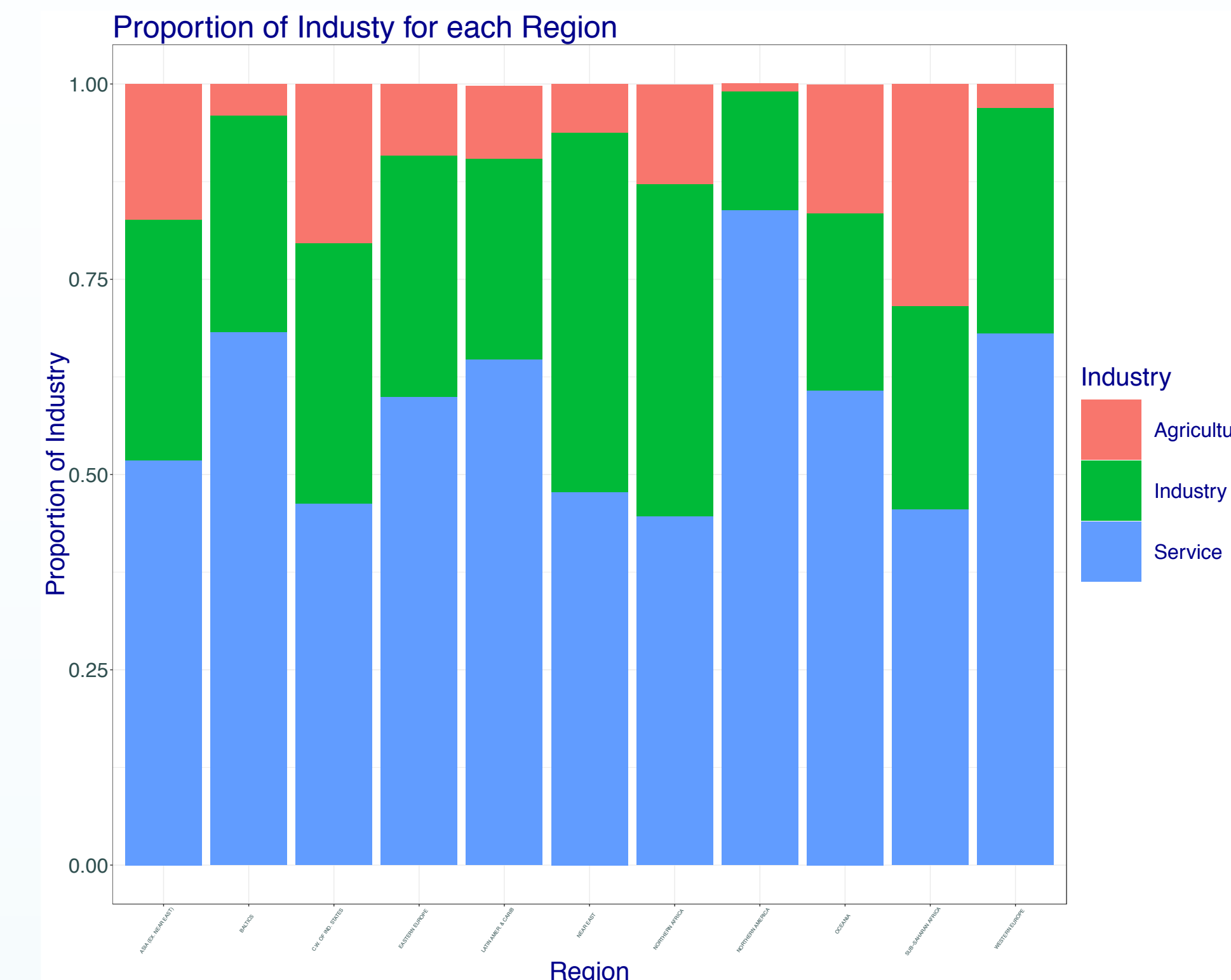


- There is a positive polynomial relationship between GDP per capita and literacy rate.
- Regions like Sub-Saharan Africa have lower literacy rate and lower log GDP per capita.
- Regions like Western Europe have higher literacy rate and higher log GDP per capita.

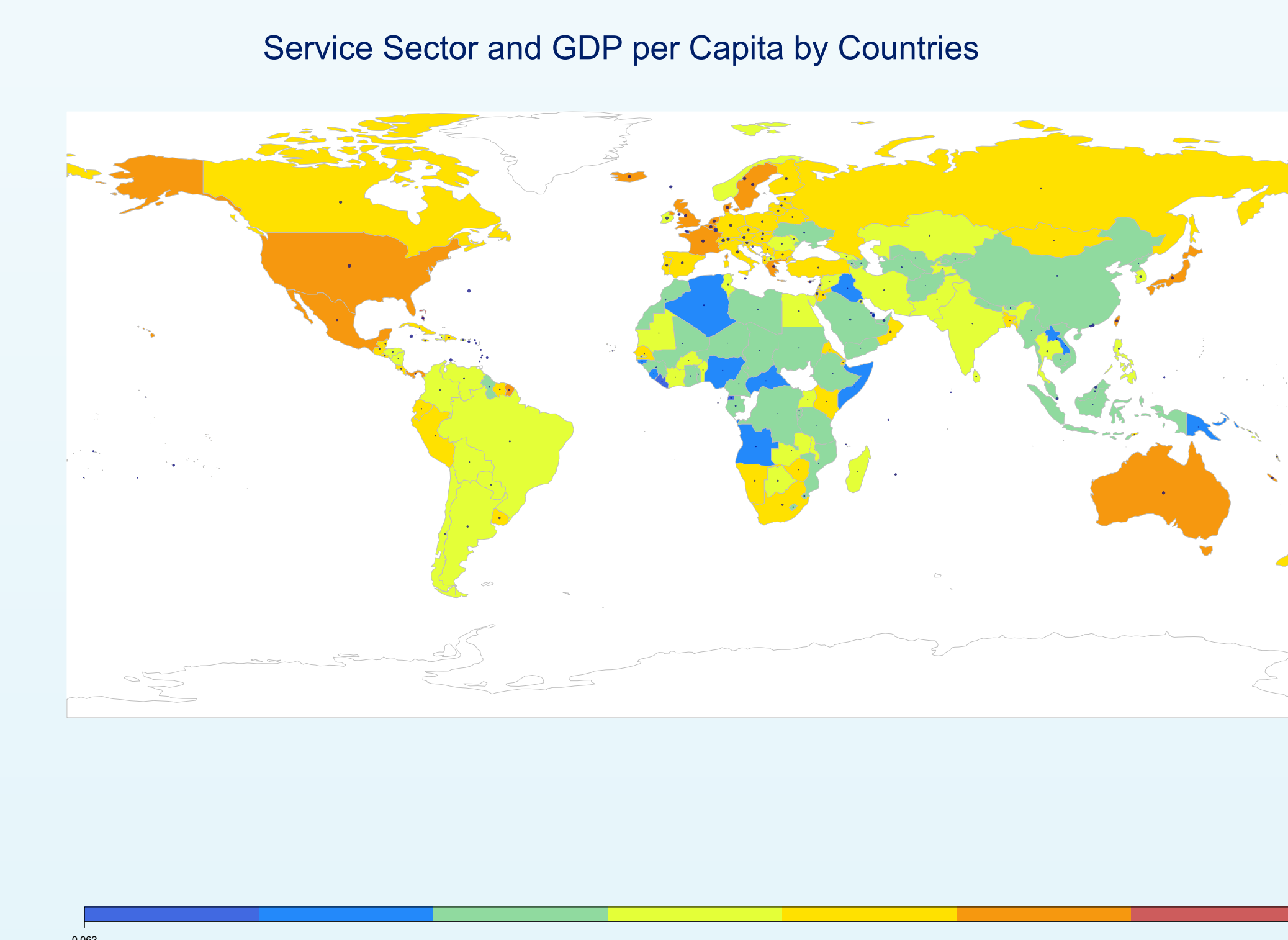


- Regions such as North America, Asia, Europe, and Australia have much higher literacy rates compared to countries in Africa and Southeast Asia.
- This relates to the scatter plot as these regions encompass countries that have higher GDP per capita.

Service Sector

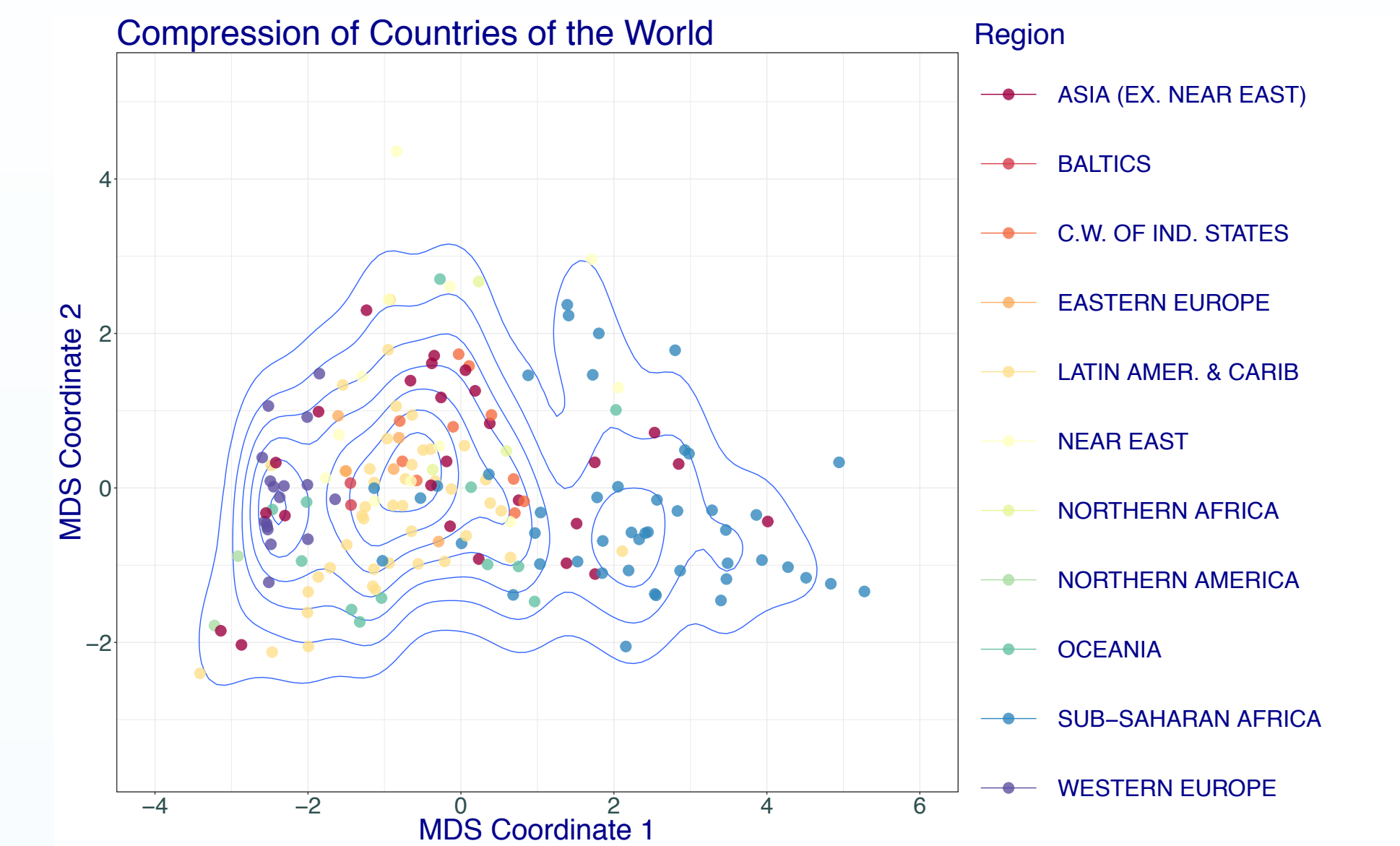


- Service sector contributes most to GDP per capita among regions.
- Northern America, Western Europe and Baltics have relatively higher service sector among regions.



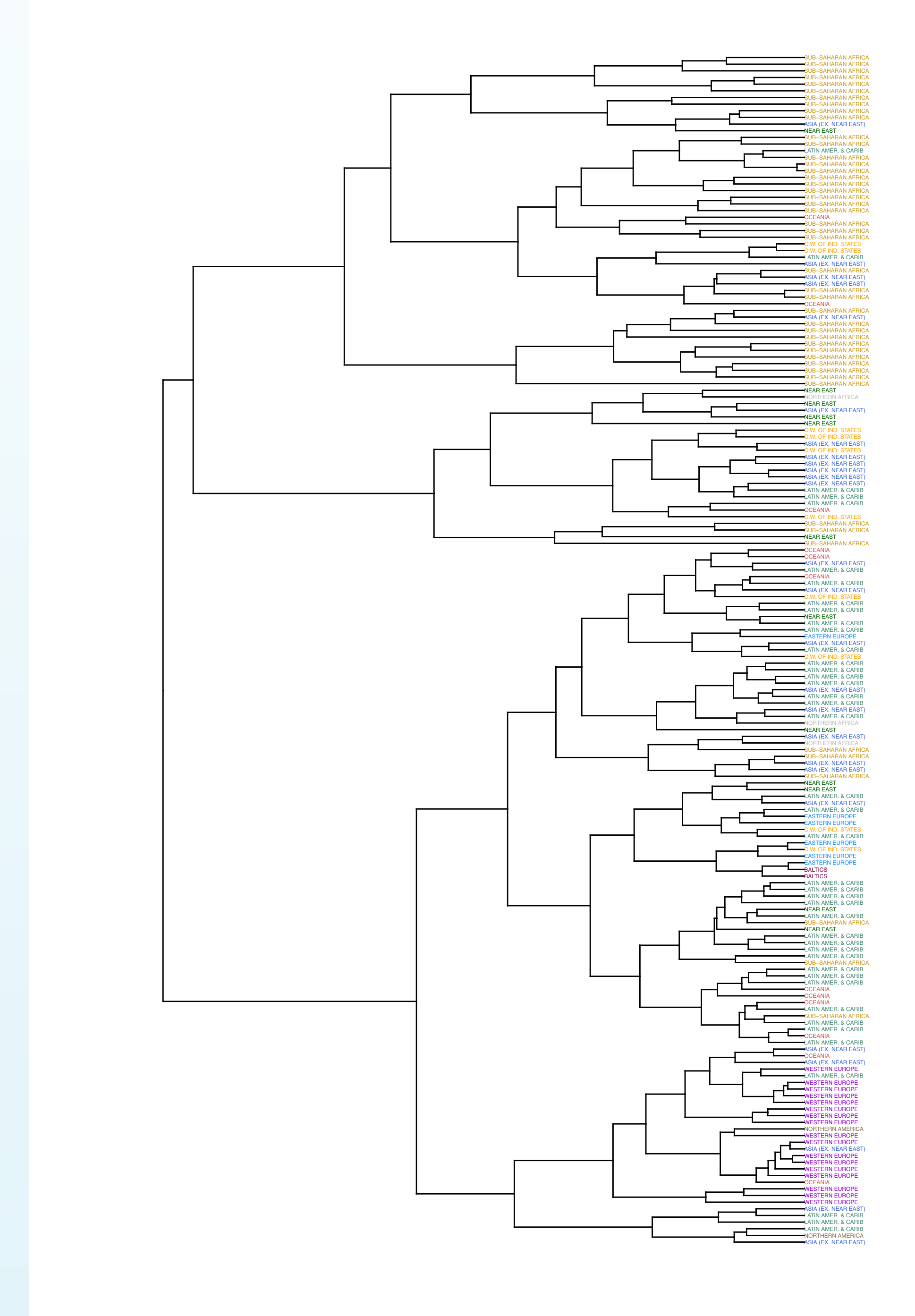
- Countries that have higher service sectors tend to have higher GDP per capita.
- We can confirm that higher GDP per capita regions like Northern America, Western Europe and Baltics have higher service sector.

Clusters of Regions



- Density plot based on variables birth rate, population, industry, literacy rate and GDP per capita.
- There are three main clusters, each corresponding to specific regions: Western European, Latin America and Caribbean, and Sub-Saharan Africa.

Cluster Dendrogram of Countries Regions



- Western European, Latin America and Sub-Saharan Africa make up the major clusters.
- Some smaller clusters of regions are Asia, Eastern Europe and Oceania.